The Dissertation Committee for Mi Jang certifies that this is the approved version of the following dissertation:

PROSODICALLY DRIVEN PHONETIC PROPERTIES IN THE PRODUCTION AND PERCEPTION OF SPOKEN KOREAN

Committee:

Harvey M. Sussman, Co-Supervisor

Randy L. Diehl, Co-Supervisor

Megan Crowhurst

Richard P. Meier

Anthony C. Woodbury

PROSODICALLY DRIVEN PHONETIC PROPERTIES IN THE PRODUCTION AND PERCEPTION OF SPOKEN KOREAN

by

Mi Jang, B.A.; M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

In Partial Fulfillments

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

August 2009

Dedication

To my husband and daughter,

with gratitude and love

Acknowledgements

For every step towards completing this dissertation, I am much indebted to the many people who have surrounded me and championed my work. I would like to take this opportunity to acknowledge each of their contributions and support.

First of all, my deepest appreciation goes to my two supervisors, Dr. Randy Diehl and Dr. Harvey Sussman. Without their unflagging support and consistent encouragement, I would not have had this precious opportunity to express my most sincere gratitude to them in these acknowledgements. I am as much indebted to Dr. Harvey Sussman, one of my supervisors. Two and half years ago when I was in deep despair, he generously reached out with warmth and took me by the hand. His thought provoking suggestions and encouragements helped me in all the time of research and writing of this dissertation. He guided me with patience toward the completion of this dissertation by sharing his time and opinions whenever I needed his advice.

If the fruit of this study is more than what it was originally intended to bear, it is mainly due to my other supervisor, Dr. Randy Diehl. His Speech Perception class was one of the most influential courses I have ever taken and introduced me to the world of speech perception. Although he is very busy as the Dean of the College of Liberal Arts, he has always been within arm's length when guidance, advice, and care were highly needed.

Needless to say, I wish to express my gratitude to my other committee members, Dr. Richard Meier, Dr. Anthony Woodbury, and Dr. Megan Crowhurst. Special thanks should go to Dr. Megan Crowhurst and Dr. Scott Myers who were the supervisors of my qualifying paper. They broadened my knowledge of phonology and phonetics. The experience of working as a teaching assistant for each of them kept me fascinated with the world of sounds.

My thanks go to the anonymous subjects who gave their precious time to participate in the experiments. I want to thank Dr. Greg Hixon in Psychology Department and ITS research consultants for their prompt and careful consultations on my statistical analyses.

I would like to express my sincere gratitude to my professors in Korea who were great mentors and inspired me to continue my studies: Professor Jae-Young Lee at Seoul National University and Professor Hyang-Sook Sohn at Kyungpook National University. I wish to deeply thank the professors who guided me to other sub-fields of linguistics at Seoul National University: Professor Ki-Sun Hong, Professor Mi-Jeong Song, and Professor Eun-Jung Yoo.

I feel so lucky that many wonderful people surrounded me, providing care and warm encouragement. I want to express my special thanks to my lovable friends, Cynthia Hansen, Taryne Hallet, and Robin Fletcher for giving me their friendship, as deep and as rich as friendship can be. I also extend my thanks to Jess White-Sustaita, Kate Shaw-Points, Sang-Hoon Park, Tae-Sun Moon and Yoon-Jung Bae for the support they have lent me during my stay in Austin.

I would like to thank my family. My deepest appreciation goes to my grandmother and late grandfather who have prayed for my academic success and good health countless days and nights. When my late grandfather was alive, he asked me to bring the certificate of Ph. D to his tomb if he were to pass away before I finished. I will finally fulfill this promise.

Finally, and most importantly, I would like to thank my husband, Chang-Hee Jung. His support, quiet patience and unwavering love were undeniably the bedrock upon which the past six years of my life have been built. With pleasure, my special thanks goes to my daughter, Jacqueline Seunghyun Jung for being in the world. Nothing can be comparable to you. Thank you for being such a joy in my life.

Prosodically Driven Phonetic Properties

in the Production and Perception of Spoken Korean

Mi Jang, Ph.D.

The University of Texas at Austin, 2009

Co-supervisors: Harvey M. Sussman and Randy L. Diehl

The focus of this study was to explore how prosodic position and word type affect the phonetic structure and resulting perceptual identification of Korean stops and fricatives. When there is less contextual information, speakers tend to produce clearer speech. For example, consonants at the beginning of prosodic domains, such as syllables, words or phrases, are known to be more clearly articulated and distinguishable than later-occurring consonants. However, it is not yet clear whether the prosodically conditioned realizations of a segment are perceptually distinctive in continuous speech. In addition, there are few studies examining whether the properties of prosodic domain-initial segments are affected by the information content of words (real vs. nonsense words).

The acoustic properties of stops and fricatives were compared across IP, PP and Wd-initial positions both in real and nonsense words. It was found that segments in the higher prosodic domain-initial positions showed enhanced durational properties

compared to the lower prosodic domain-initial positions. However, the enhancing strategies were different among phonation types. Relative to lenis and aspirated stops, and lenis fricatives, tense stops and fricatives showed less consistent variation as a function of prosodic position and word type.

In the perception study, the identification error rates and reaction time for same-spliced CVs were compared to those for cross-spliced CVs. Korean listeners identified the same-spliced CVs more accurately and faster than cross-spliced CVs. In addition, the distinctive acoustic properties of each prosodic domain-initial position were perceptually distinguished by Korean listeners. Due to relatively shorter duration and less distinctive contrast, the target CVs extracted from lower prosodic domain-initial positions caused more confusion in the identification of target segments.

In conclusion, this study provides the evidence that speakers modulate their speech clarity depending on information content. By enhancing phonetic properties and phonological contrast, speakers tend to provide perceptual cues for prosodic positions with less contextual information.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

The focus of this research was to explore how prosodic position and word type affect the phonetic structure and resulting perceptual identification of Korean stops and fricatives. Speakers try to achieve sufficient distinctiveness in articulating segments to increase intelligibility for listeners (Lindblom (1990)). Utterance clarity has been found to vary with information content (Hay et al. (2006)). When there is less contextual information (with higher information content), talkers tend to produce clear speech. For example, consonants at the beginning of prosodic domains, such as syllables, words or phrases, are more clearly articulated and distinguishable than later-occurring consonants (Browman and Goldstain (1995), Redford and Diehl (1999), Keating et al. (1999), Cho and Keating (2001)). By enhancing phonetic properties and phonological contrast, talkers tend to provide perceptual cues for the prosodic position with higher information content.

Segmental properties are known to be affected by prosodic structure. Acoustic and articulatory studies have focused on both phrase-final intonational contrasts and final lengthening (Pierrehumbert & Beckman (1988), Shattuck-Hufnagel et al. (1996)) as well as initial prosodic domains across syllable, word and phrasal levels (Pirrehumbert and Talkin (1992), Browman and Goldstein (1995), Redford and Diehl (1999), Keating et al. (1999), Fougeron (2001)). It has been recently found that the initial segment of prosodic domains is more strongly articulated and longer relative to prosodic domain-medial segments (Fougeron & Keating (1997), Cho & Keating (2001), Keating et al. (2003)), and there is reduced coarticulation between phonemes that span the phonological phrase boundary (Byrd et al. (2000), Cho (2004)).

Several studies examining word edges have shown that articulations of the tongue, lips, velum and glottis differ in magnitude in word-initial versus non-initial position (Browman and Goldstein (1992)). Initial consonants had a higher velum position for both nasal (Krakow (1989)) and oral segments (Vaissiére (1988)). In studies of vowel formant transitions at the syllable level, F2 trajectories were more distinctive for place value in CV than in VC syllables (Sussman et al. (1997)). Initial segments at the syllable and word levels were found to be resistant to reduction or lenition processes regarding synchronic and diachronic variants, relative to medial or final segments (Bell & Hooper (1978), Ohala & Kawasaki (1984)). The realization of both /h/ and /ʔ/ in English was influenced by prosodic position as both phrasal accent and phrasal boundary increased the magnitude and duration of the glottal gestures for the segments (Pierrehumbert & Talkin (1992)).

With regard to linguopalatal articulation at the word level, initial consonants were found to have a greater linguopalatal constriction than medial consonants (Byrd (1994), Keating et al. (1999)). At phrase and sentence levels, more linguopalatal contact for coronal stops was shown in initial position of higher domain than in lower ones, as measured by electropalatography (e.g. in English (Keating et al. (1999)), in French

(Fougeron and Keating (1997), Fougeron (2001)) and in Korean (Cho & Keating (2001)). Fougeron & Keating (1997) described the increase in linguopalatal contact in domain-initial position as articulatory strengthening, meaning the articulation of a consonant is more extreme in initial position compared to medial, and more extreme at the beginning of higher level constituents than at the beginning of lower ones. The articulatory properties found in initial segments reflect the hierarchical organization of the constituents.

In addition to the effect of prosodic domain-initial position, different word type is expected to influence speech clarity. Since nonsense word condition yields higher information content relative to real word condition, phonetic segments should be more clearly articulated in nonsense words than in real words. Analogous studies were conducted related with frequency of occurrence. For example, more frequent items tend to be produced with more quickly and with greater coarticulation than comparable items of lower frequency of occurrence since lower frequency of occurrence yields higher information content (Wright (2003), Pluymaekers et al. (2005)). There are few studies comparing the segmental properties between real and nonsense words but it is worth examining how phonetic properties of segments in the prosodic domain-initial positions vary as a function of different word type.

Although the enhanced acoustic properties of prosodic domain-initial segments have been studied in several languages, it is not yet clear whether listeners are sensitive to the prosodically conditioned realizations of a segment in the perception of continuous speech. The current study therefore attempted to examine whether the prosodically driven phonetic properties of the segment affect the perception of the segment in continuous Korean speech. In addition, since Korean has a unique three-way contrast in stops and a two-way contrast in fricatives, it is also worth examining how phonological contrast is enhanced according to prosodic position and word type and whether the enhanced distinction is also reflected in speech perception.

The first part of this chapter provides a short summary of previous studies on Korean stops and fricatives, and the motivation and organization of the study are presented in the remainder of the chapter.

1.1 Korean Stops and Fricatives

Korean has heavily aspirated, lightly aspirated, and laryngealized unaspirated stops in utterance-initial position. The three different categories are referred to as lenis ([p, t, k]), aspirated ([p^h, t^h, k^h]) and fortis ([p', t', k'])[1] and each of these occurs at three places of articulation: bilabial, alveolar and velar. They are all voiceless and use the pulmonic egressive airstream. However, there is only a two-way contrast in dental fricatives: lenis /s/ and fortis /s'/. The inventory of Korean stops and fricatives is given in Table 1. The minimal contrasts of Korean stops and fricatives are shown in Table 2.

---

[1] /t'/ represents the fortis stop, for which there is no official IPA transcription.

Table 1.1 Korean stops and fricatives: lenis, fortis and/or aspirated distinctions

Lenis	p	t	k	ʧ	s
Aspirated	p^h	t^h	k^h	$ʧ^h$	
Fortis	p’	t’	k’	ʧ’	s’

Table 1.2 Minimal contrasts for Korean stops and fricatives in word-initial position (Cho et al. (2002))

Lenis	Aspirated	Fortis
paŋ ‘room’	p^h aŋ ‘bang’	p’ aŋ ‘bread’
tal ‘moon’	t^hal ‘mask’	t’al ‘daughter’
keta ‘to fold up’	k^heta ‘to dig up’	k’eta ‘to break’
sata ‘to buy’		s’ata ‘to wrap’

The phonetic correlates of this three-way contrast have been examined by many researchers. In Lisker and Abramson (1964), VOT was a highly effective phonetic cue for differentiating stop categories with different laryngeal categories in a variety of languages. They found that Korean fortis stops are likely to have almost zero VOT, aspirated ones the longest VOT and lenis ones an intermediate VOT. However, several researchers noted that the VOT ranges overlap in Korean stops (Kim (1965), Han & Weitzman (1970), Cho et al. (2002)). In perceptual studies, it was also found that the difference in VOT alone was not sufficient for distinguishing the Korean three-way phonation contrast (Han and Weitzman (1970), Abramson & Lisker (1972), Han (1996), Kim, Beddor, & Horrocks (2002)).

Han and Weitzman (1970) reported that intensity build-up in a vocalic segment was highest after fortis consonants, intermediate after aspirated consonants, and lowest after lenis ones. It has also been noted that F0 after fortis and aspirated consonants is higher than that of lenis ones (Kim (1965), Han and Weitzman (1970), Hardcastle (1973), Kagaya (1974), Cho et al (2002)).

Results from aerodynamic studies in Dart (1987) and Cho et al. (2002) showed that maximum oral airflow rates were highest for the aspirated, lowest for the fortis, and intermediate for the lenis consonants word-initially. In the study of intraoral air pressure in Cho et al. (2002), it was highest for the aspirated, lowest for the lenis and intermediate for the fortis stops but there was no significant difference between aspirated and fortis stops.

Several fiberscoptic studies of Korean (Kagaya (1974), Jun, Beckman and Lee (1998)) have reported that glottal opening was smallest for fortis stops, moderate for lenis stops and largest for aspirated stops. In stroboscopic-cine MRI experiments, Kim et al. (2005) found that the glottis opened much wider for the aspirated consonants than for the fortis and lenis ones word-medially and word-initially and that the movement of the tongue blade and closure duration varied from short to long in the order lenis < aspirated < fortis.

The voice quality of the onset of the following vowel was shown to be influenced by the preceding consonant (Abberton (1972), Ahn (1999), Cho et al. (2002)). The onset of vowels after lenis stops had a breathy voice as indicated by positive [H1-H2] values (the difference in amplitude between the first and second harmonics), whereas fortis stops had some of the characteristics of creaky voice with smaller or negative [H1-H2] values (Abberton (1972)). Cho et al. (2002) examined [H1-H2] values for vowel onset after three different Korean stops in two dialects (Seoul and Cheju Korean). The [H1-H2] values of vowel onset were greatest for lenis stops, intermediate for aspirated stops, and smallest for fortis stops in the Seoul dialect, but [H1-H2] of the aspirated stop was close to that of the lenis stops for some speakers, or to that of the fortis stop for some others in the Cheju dialect. However, Ahn (1999) found that both raw and normalized [H1-H2] values were greatest for the aspirated stops, intermediate for the lenis stops, and smallest for fortis stops in Seoul Korean. It seems that the results from [H1-H2] values are inconsistent across speakers and across dialects. Cho et al. (2002) interpreted that the difference could be caused by the relatively younger age group of Ahn (1999)'s speakers and that procedural differences may have caused the inconsistency between the two studies. Table 3 summarizes the phonetic correlates across the Korean stop categories in the previous studies.

Table 1.3 Summary of phonetic correlates of Korean stop categories

Acoustic properties	Order in stop category
VOT	aspirated >> lenis >> fortis
Glottal opening	aspirated >> lenis >> fortis
Intensity build-up	fortis >> aspirated>> lenis
F0	fortis, aspirated >> lenis
Maximum oral airflow	aspirated > lenis >> fortis
Intraoral air pressure	aspirated >> fortis >> lenis
H1-H2	aspirated, lenis >> fortis

Unlike stops, there is only a two-way contrast between fricatives, namely the lenis fricative, /s/[2] and fortis fricative, /s'/ in Korean. In comparing fricative durations, there are discrepancies across previous studies. In Yoon (1998), /s'/ had longer duration than /s/ in non-high vowel contexts. Park (1999) reported that there was no significant durational difference between /s/ and /s'/ in a VCV context. Cho et al. (2002) showed that /s/ consists of two components (frication and aspiration) and that the duration of /s/ is significantly longer than that of /s'/ in word-initial position. However, if the aspiration

[2] The categorization of /s/ has been controversial in previous research. In Kagaya (1974) and Jun, Beckman, & Lee (1998), /s/ was reported to have a glottal opening configuration similar to aspirated stops in fiberscopic data. Park (1999) claimed that /s/ should be categorized in the aspirated category because the vowel onset after /s/ is breathier than after /s'/. On the contrary, Cho et al.(2002) suggested that /s/ is regarded as a lenis segment because /s/ has a similar breathy voice quality to vowels after the lenis stops and F0 after /s/ is lower than that after an aspirated stop. In addition, they reported that half the tokens of /s/ were fully voiced intervocalically, which was not shown in the aspirated stop category. In some phonological processes such as Post-Obstruent Tensing (S. Kim (2001)), /s/ becomes tense following an obstruent just as the lenis stops do. Since this study is not concerned with the categorization of /s/, /s/ will be referred to as lenis fricative.

portion is excluded, /s'/ was reported to be longer than /s/ in Seoul Korean in Cho et al. (2002).

With regard to centroid frequency[3], Cho et al. (2002) claimed that /s'/ was produced with a relatively smaller front cavity because the centroid frequency was higher for /s'/ than /s/ in Seoul Korean. However, half of the Cheju speakers did not show this distinction. Park (1999) and Yoon (1998) did not find a significant difference in spectral peak between the two fricatives.

In the results of Seoul Korean in Cho et al. (2002), there was no difference in F0 between the two fricatives in Seoul Korean, but [H1-H2] was significantly higher for /s/ (positive) than for /s'/(negative), indicating the breathiness of vowels after /s/ and creakiness of vowels after /s'/. The summary of phonetic correlates across fricative pairs is shown in Table 4.

Table 1.4 Summary of phonetic correlates of Korean fricative categories

Acoustic properties	Order in fricative category
Fricative duration	lenis >> fortis
Centroid frequency	fortis >> lenis
H1-H2	lenis >> fortis
F0	no difference between them

[3] The centroid was the center of gravity of a defined part of the spectrum, each frequency being weighted according to its amplitude (Cho et al. (2002)).

1.2 Prosodic Hierarchy

In order to compare the phonetic properties of different prosodic domains, the target segments in the previous studies were coded according to their position within prosodic hierarchy theory (Nespor & Vogel (1986), Selkirk (1986), Beckman & Pierrehumbert (1986)). Speech utterances are hierarchically organized, with higher units being decomposed into lower constituents (Nespor & Vogel (1986), Selkirk (1986)). The prosodic units are marked by suprasegmental features such as intonational events and/or final lengthening. A number of studies have shown that prosodic grouping of an utterance is not isomorphic to syntactic structure[4], and the organizational structure of underlying spoken utterances corresponds to prosodic rather than syntactic hierarchies. Prosodic constituents have been defined within several different theoretical frameworks, and in terms of the domains of: (1) phonological rules, (2) intonation, and (3) rhythmic prominence (Shattuck-Hufnagel & Turk (1996:209)). Since the focus of this study is not concerned with a particular constituent in the prosodic hierarchy, I will follow the general hierarchical structure of spoken utterances, without relying on the detailed assumptions of a particular theoretical proposal.

[4] It has been noted that there are many discrepancies between the syntactic structure of a sentence and the prosody of an utterance of that sentence, despite a general correspondence between the two sets of structures (Selkirk (1986), Nespor & Vogel (1986)). Prosodic phrasing is also affected by non-syntactic factors such as speaking rate and constituent length. Both extrasyntactic and syntactic factors influence the speaker's choice of prosodic shape for an utterance.

An Intonational Phrase (IP) is defined by a complete intonational contour, including a final boundary tone (Shattuck-Hufnagel & Turk (1996)). It consists of a whole clause or sentence and is often marked by a pause with final lengthening. A Phonological phrase boundary coincides with the syntactic phrase boundary and is characterized by final lengthening and a single pitch contour. The Phonological Phrase (PP) has been defined in syntactic terms, although precise syntactic definitions differ between theorists. Under Selkirk's end-based theory (Selkirk (1986)), a PP in Korean is defined as {left, X^{MAX}} where the PP boundary exists at the left end of a maximal projection such as Noun Phrase, Verb Phrase, etc. In terms of the relation-based theory (Nespor & Vogel (1986)), the PP in Korean is formed by adjoining the head and its adjacent complement in the maximal projection. On the other hand, this phrase also appears to be constrained by non-syntactic factors such as pitch accent and phrasal tone (Jun (1993, 1998)) which is termed as Accentual Phrase (AP). An AP usually consists of content words and function words with an associated phrasal tone pattern. The Korean AP is known to have the underlying phrasal tone sequences, LHLH or HHLH[5]. The two models are shown in Fig. 1, which gives sample structures of the hierarchical organization of the prosodic domains. The syntactic-based model and intonation-based model are shown in Fig.1.a and Fig.1.b, respectively. In this study, the intermediate

---

[5] Jun (2000) presented that an AP in Seoul Korean is mostly marked by LHLH or HHLH when the AP is longer than three syllables. The AP initial tone in Seoul Korean is marked with H when the segment is aspirated or tense, but L otherwise. But it was noted that an AP can be realized in at least fourteen different tonal patterns, with more variation when the AP has fewer than three syllables (i.e., LH, LHH, LLH, LHLH, HH, HLH, HHLH, LL, HL, LHL, HHL, HLL, LHLL, HHLL). She mentioned that those patterns are neither distinctive in meaning nor predictable.

phrase between IP and Prosodic Word is formed by considering both syntactic and intonation-based approaches and will be referred to as PP. A Prosodic Word (Wd) contains one lexical head, grouped with some functional elements (i.e., case or postpositional marker).

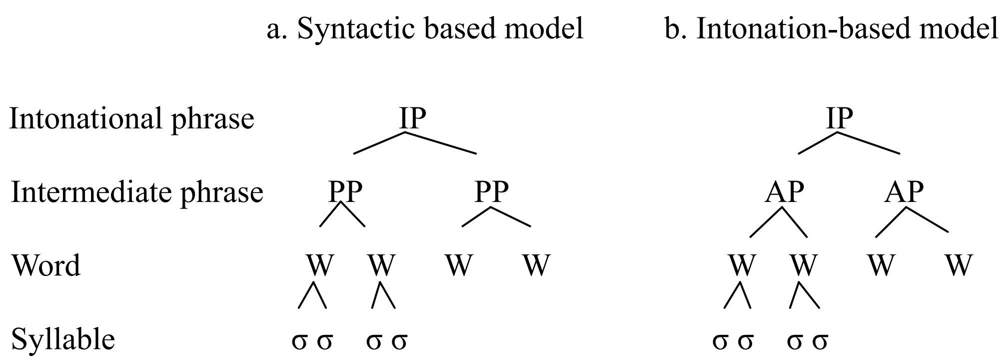


Figure 1.1 Prosodic structure of Korean

1.3 Motivation of the Study

1.3.1 Motivation for Studying Korean

Korean is one of the languages that show prosodically conditioned properties on the initial segments of prosodic domains. Both phonological and/or phonetic processes are sensitive to prosodic domains. It was noted that there are some phonological and/or phonetic rules that apply only in certain prosodic domains. Jun (1993, 1998) proposed that Accentual Phrase is the domain of application of Post Obstruent Tensing and Lenis Stop Voicing rules[6] in Korean. That is, the voicing or tensing of a consonant is constrained by the consonant's position in a phrasal domain. Jang (2008) reported that the Aspiration Merger rule is sensitive to phonological phrasing and occurs within the Phonological Phrase.

1.3.1.1 Stops and Prosodic Domains

Cho & Keating (2001) examined the effect of prosodic position on the segmental properties of the Korean consonants /n, t, t^h, t'/ in initial position in five prosodic domains of the Korean prosodic hierarchy. A progressively increasing trend was found in domain-initial lengthening of closure duration, VOT, and total voiceless interval (Cho & Jun (2000), Cho & Keating (2001)). The results of VOT in this study showed the pattern Ui[7] > IPi> APi > Wi. With regard to the increasing trend in phrase-initial position, Jang (2008) also reported that VOT of Korean aspirated /p^h/ was significantly longer at the beginning of a Phonological phrase (PP) than in the medial position of a PP. Acoustic total voiceless interval combined voiceless closure duration and VOT, indexing the

[6] Post Obstruent Tensing is a process that makes onset lenis consonants tensed when they follow an obstruent coda (e.g., [ʧip] 'house' + [pak] 'outside' → [ʧipp'ak] 'outside the house'). The Lenis Stop Voicing rule shows that underlying voiceless lenis stops become voiced in intervocalic position (e.g., [pabo] 'fool', [kɯ babo] 'the fool').

[7] The lowercase i represents the initial position of each prosodic domain. For example, Ui and APi indicate Utterance-initial position and Accentual Phrase-initial position, respectively.

duration of glottal opening. Since aspirated stops showed the pattern IPi > APi> Wi for total voiceless interval, it revealed that the duration of glottal opening is greater in the higher prosodic positions than in lower prosodic positions. In the percentage of voicing of the entire stop closure duration, a progressively decreasing pattern was found, namely IPi < APi < Wi. Since the peak linguopalatal contact showed a progressively increasing pattern Ui > IPi> APi > Wi, they claimed that there is a strong correlation of linguopalatal contact with duration, articulatorily and acoustically.

In comparing RMS burst energy to each domain-initial vowel, Cho & Keating (2001) reported that RMS burst energy was smaller for Ui and IPi than APi and Wi for /t^h, t'/ but that there was no systematic pattern for /t/. But the results were not statistically significant. Since Cho & Keating (2001) did not normalize the RMS burst energy across speakers, it seems that the effect of prosodic domain-initial position could have been washed out by individual differences. In order to compare energy values of each speaker, RMS burst energy should be normalized.

It was noted that RMS burst energy is dependent upon articulatory/aerodynamic characteristics of the stop release gesture and higher intraoral pressure build-up during the stop closure may give rise to higher RMS burst energy (Cho et al. (2002), Cho & Keating (2001, 2007). For example, when the percentage value of the burst energy relative to the energy at the midpoint of the vowel was compared across three Korean stops, aspirated stops were shown to have greater relative RMS burst energy than the other stop categories due to greater airflow and high air pressure (Cho et al. (2002)).

Cho & Keating (2007)[8] showed that the RMS burst energy of /t/ in English was lower in Utterance-initial position than in Utterance-medial position. They explained that the overall low level of intraoral pressure is possible factors for the low burst energy in Utterance-initial position. Since the effect of prosodic domain-initial position on RMS burst energy found in the previous studies of Korean is inconclusive, it is necessary to examine the difference for normalized RMS burst energy across different prosodic domain-initial positions. When the target stops are placed in higher prosodic domain-initial positions, they are expected to have lower RMS burst energy.

The prosodic domain-initial effect on vowel following stops has not been studied. It has been found that Korean stop and fricative categories are distinguished in part by [H1-H2] values and F0 (Cho et al. (2002)). In a pilot study of prosodic position effect on [H1-H2], it was found that [H1-H2] values after stops were greater in lower prosodic domain-initial positions than in higher prosodic domain-initial positions (that is, the creakiness of the vowel is less). When it comes to F0, there is substantial literature documenting that fundamental frequency drops gradually from the beginning to the end of spoken sentences or phrases (Umeda (1982), Thorsen (1985)) although the declination of F0 is known to be speech style-dependent. In the prosodic hierarchy, a constituent at

[8] It was also noted that the larger linguopalatal contact for /t/ may induce longer release duration, resulting in reduced peak burst energy (Stevens et al. (1986).

one level is thought to be composed exclusively of one or more constituents from the next level down, following the Strict layer Hypothesis (Selkirk (1984)). For example, an IP is composed of PPs, and a PP is composed of Wds. So, F0 after target segments is expected to be higher in the higher prosodic domain-initial positions than in lower prosodic-initial positions.

With regard to phrase final lengthening, the duration of the final vowel of the prosodic domain was found to be longest for either U-final or IP-final, intermediate for AP-final and shortest for W-final in Cho & Keating (2001). It is worth examining phrase final lengthening in addition to the prosodic domain-initial effects because they both provide strong evidence for a prosodic domain boundary. Either or both properties may help listeners to perceive the prosodic domain boundary in speech. Table 5 summarizes the results of the previous studies on acoustic and articulatory variation of Korean stops in the initial position of different prosodic levels.

Table 1.5 Summary of studies showing acoustic and articulatory variation of Korean stops in different prosodic initial positions

Acoustic or articulatory properties	Results
Peak linguopalatal contact	Ui >> IPi>> APi>>Wi
Seal duration[9]	Ui >> IPi>> APi>>Wi
VOT	Ui >> IPi>> APi>>Wi
Total voiceless interval for $/t^h/$	IPi >> APi>> Wi
Percentage of voicing of the entire stop closure duration	Wi >> APi>> IPi
RMS burst energy	APi, Wi > Ui, IPi[10]

1.3.1.2 Fricatives and Prosodic Domains

In comparison to stops, there are only a few studies focusing on the domain-initial properties of fricatives. In a study of French consonants, Fougeron (2001) showed that the domain-initial strengthening effect is less robust for the fricative /s/ than for the stop /t/ in terms of linguopalatal contact. For linguopalatal contact of /s/, it was found that /s/ is less systematically affected by prosodic position compared to the other consonants. In addition, it was noted that /s/ showed fewer positional differences in the comparison

9 Seal duration represents the duration between the first and the last frame in which the oral cavity was completely sealed (Cho & Keating (2001): pp 161).

10 Cho & Keating (2001) excluded /t/ in the results of RMS burst energy due to the presence of voicing throughout the consonant in word-initial position.

between onset and coda in English (Byrd (1994)).

In a study of Korean fricatives, S. Kim (2001) examined the acoustic and articulatory properties in two divided sets of domains - higher prosodic domains (IPi, APi, and APm (AP-medial position)) and lower prosodic domains (Wi and Si (syllable-initial position)). In degree of linguopalatal contact[11], it was found that both fricatives showed difference between IP and AP but no difference between APi and APm. In the results of acoustic properties, fricative duration was longer, and centroid frequency was higher in higher prosodic domains than in lower prosodic domains. [H1-H2][12] value for fortis fricative, /s'/ was lower in higher domains than in lower domains but there was no significant difference across different prosodic levels for /s/. However, S. Kim (2001) argued that the prosodic domain effect on Korean fricatives was not as strong as the one on Korean stops. Furthermore, the two speakers participating in the study did not show the same results. For example, in the results for fricative duration, both speakers showed significant differences among three prosodic domains for /s/, but for /s'/, only one of the two speakers showed significant differences among the domains. In the results examining centroid frequency,[13] /s'/ in IP and AP initial position showed a higher centroid frequency relative to AP medial position but only one of the two speakers showed a

[11] Kim (2003a) measured five regions (whole, front, mid, back and channel) for linguopalatal contact but I will only report the results of the whole region.

[12] Amplitude difference between the first (H1) and second (H2) harmonics were used to distinguish between breathiness and pressed voicing quality of the vowel after different phonation types.

[13] Cho et al.(2002) compared the centroid frequency of /s/ to that of /s'/ and found out that /s'/ has a higher centroid frequency than /s/, suggesting that /s'/ is produced with a relatively smaller front cavity. They also found dialectal difference in the results of centroid frequency in that Seoul speakers produce /s/ and /s'/ with a higher centroid frequency than Cheju speakers.

significant domain difference for /s'/. Since this study was conducted with only two speakers and the results were not consistent between speakers, it is hard to conclude that the articulatory and acoustic properties of Korean fricatives are affected by prosodic structure. To find out the different prosodic domain-initial effects between stops and fricatives, and across phonation types, more detailed quantitative studies need to follow.

Fricatives are known to be less subject to articulatory variation in general because they tend to be more constrained in their articulatory and acoustic properties (Fougeron (2001)). The less articulatory variation would cause fewer prosodically dependent variations in the production and perceptual confusions of fricatives in identification tests. The summary of the previous studies showing acoustic and articulatory variation of Korean fricatives in different prosodic domain-initial positions is shown in table 6.

Table 1.6 Summary of studies showing acoustic and articulatory variation of Korean fricatives in different prosodic initial positions

Acoustic or articulatory properteis	Results
Linguopalatal contact	IPi >> APi, APm[14]
Fricative duration	IPi>> APi>>APm
Centroid frequency	IPi, APi >>APm
H1-H2	APm >> IPi, APi for /s'/

[14] APm indicates the medial position of Accentual phrase.

1.3.2 The Effect of Prosodically Conditioned Phonetic Variation on Speech Perception

The enhanced articulatory properties of the initial segment at the syllable, word or phrase level seem to be generated by speakers aiming for sufficient segmental distinctiveness in initial position. Since word-initial phonemes have been considered to be salient in lexical access and word recognition, word onsets play a critical role in determining which words are accessed (Cho et al. (2007)). They may also provide perceptual cues for the prosodic structure and information groupings by maximizing contrast in acoustic and articulatory dimensions.

In Dutch, Quené (1992) claimed that a word-initial consonant is longer than a word-final consonant and that the lengthening of the initial consonant can serve as a perceptual cue to Dutch word boundaries. Christophe et al. (2004) investigated the role in spoken-word recognition of a prosodic boundary larger than that of the word. By using word monitoring and phoneme detection tasks in French, they found evidence that lexical access for monosyllabic words was faster when the target words appeared in two-word sequences which spanned Phonological Phrase boundaries than when the sequences occurred within Phonological Phrases. With identity-priming experiments, Cho et al. (2007) showed that when a two-word sequence straddled a Wd-boundary, the strengthened initial CV of the post-boundary word (i.e., CVs from IP-initial position) assisted listeners most in the recognition of pre-boundary words. When the two-word sequence is straddling a Wd boundary or IP boundary in a carrier sentence, the recognition of pre-boundary words were found to be better at an IP boundary than at a Wd boundary.

The increased articulatory or acoustic contrast between segments straddling a prosodic boundary may facilitate lexical access to the domain-initial lexical item which has less contextual or discourse information than lexical items occurring domain-medially. This contrast could contribute to marking prosodic boundary, and thus help listeners to parse the incoming speech signal into words and phrases.

In a perception study of Korean, Baker (2002) compared the error rates for the identification of mismatched CVs from Ui, IPi, and Si in the perception of Korean coronal stops. It was found that error rates of mismatched CVs were higher for lenis stops than for aspirated and fortis stops when the target CVs were taken from higher prosodic domains. The distinction between the highest levels (i.e., Ui and IPi) and lowest level (i.e., syllable) has been studied but the results did not reveal a perceptual distinction at the intermediate level. Since the intermediate level was also noted to be distinguished in the results of the previous production studies, it is necessary to find out whether the distinction of prosodic levels suggested by the production studies affects the perception of Korean utterances.

In addition, the target CVs in Baker (2002) were spliced from the onset of stop closure and the closure duration of U and IP-initial stops was inferred from the average sealed closure durations in Cho & Keating's experiment (2001). This study did not control the potential pause at U or IP boundaries of its own speaker who recorded the stimuli. Cho et al. (2007) noted that a pause associated with IP is independent of stop production and cannot be attributable to domain-initial properties. Since the identification error rates in Baker (2002) were very small (overall 93 % of correct responses for /t'/, 75 % for /t/ and 92% for /t^h/), the potential pause seems to provide information for the target prosodic boundary and thus may have washed out the prosodic domain-initial effect on perception. So, to exclude the possible effect of a pause before the higher prosodic levels, the onset of the CV needs to be defined as the release of the stop closure and the duration of the silent stop closure should not be included in the mismatched CVs.

Since there are no previous studies investigating the prosodic domain-initial properties of Korean stops and fricatives in both production and perception, I investigate how phonetic properties found in production affect the perception of the target segments.

1.3.3 Variability in the Realization of Prosodically Driven Articulatory Properties

The realization of prosodically dependent properties is quite variable. Not all the prosodic constituents studied were distinguished by articulatory variation. In English, three to four constituents out of five prosodic constituents (i.e., Utterance (U), Intonational Phrase (IP), Phonological Phrase (PP), Prosodic Word (Wd), and Syllable (S)) were distinguished by the amount of linguopalatal contact of their initial consonant,

/n/ (Fougeron & Keating (1997)). In Korean, three to five prosodic constituents were distinguished out of five prosodic constituents (by lingual articulation with segments, /t/, /t^h/, /t'/ and /n/ in Cho and Keating (2001)). In addition, not all of the subjects distinguished the prosodic constituents in the same way. Fougeron (2001) noted that the realization of the prosodically-driven articulatory properties varied depending on the segment type, the articulator, the speaker and the constituent. The most robust distinctions observed were the ones made between the most extreme constituents of the hierarchy: the highest constituent IP[15] and the lowest constituents S (syllable) or Wd (Prosodic word). The variation which is caused by segment types and speakers may affect the perception of the segments in the different prosodic domains. It is therefore interesting to establish the extent to which the distinctions of prosodic domains found in production studies can be perceived by listeners. I examine how the variation found in the production affects the perception of the prosodic constituents in a quantitative study.

1.3.4 The Perceptual Effect of Initial CVs in Real and Nonsense Words

It has been shown that speakers modulate the clarity of speech according to demands imposed by the information content of the message. For example, lower frequency of occurrence yields higher information content and it causes enhanced

[15] Previous studies do not consistently include Utterance level. So, Intonational phrase is considered the highest constituent.

speech clarity. Vowels in more frequent items were shorter in duration and less separated in F1 and F2 values than vowels in comparable less frequent items (Guion (1995), Munson and Solomon (2004)). Cooper and Paccia-Cooper (1980) found that palatalization of [d] before [j] was more likely in high-frequency than in low-frequency words. In the study of Dutch vowels, Van Coile (1987) showed that vowels occurring in function words were shorter than the same vowels occurring in content words. In the study of Dutch affixes in Pluymaekers et al. (2005), affixes were produced to have shorter realizations in a higher frequency of the carrier word.

Speakers actively adjust articulatory effort according to the perceived difficulty of intelligibility for the listener. According to Lindblom (1990), speakers have to be more careful in production when the listener has trouble understanding but they can be less careful and coarticulate more when the listeners have better conditions for understanding. Since the segments in nonsense words tend to contain higher information content, they are expected to be more clearly articulated than those in real words.

So, it is interesting to study whether the prosodic domain-initial properties are enhanced in nonsense words. If there is more enhancing of the phonetic properties in nonsense words, the results will also show a more extended effect of prosodic domain effects.

The other possible effect of nonsense words is to exclude top-down processing. With regard to identification tests using continuous speech, target words tend to be easily identified with the help of top-down semantic processing. That is, the identification of words in continuous speech tends to be informed by the knowledge of how words fit together into sentences and into discourses. In the perception study, since there is no semantic effect on the nonsense words in the middle of continuous speech, listeners are expected to perceive the target segments not with the help of meaning but with the help of enhanced phonetic properties of the segments. So, the prosodically driven phonetic properties of prosodic domain-initial segments are expected to be more distinctively identified in nonsense word condition than in real word condition.

1.4 Goals of the Study

The goal of the production experiment was to confirm whether initial stops and fricatives in higher prosodic domains have more enhanced acoustic properties than those in lower prosodic domains and to find out whether prosodic domain-initial properties are more reinforced in nonsense words than in real words. In contrast with previous studies, the current study focused on the acoustic properties of prosodic domain-initial positions. I examined how those properties found in the production study affect the perception of the target segments. Since the previous studies examining the prosodic domain-initial properties were based on only a few speakers, and the results of RMS burst energy and centroid frequency were found to be inconsistent across segments and speakers, it is necessary to confirm the prosodic domain-initial effect on the acoustic properties of Korean stops and fricatives in a quantitative study with a relatively large subject pool. I also compared voice quality ([H1-H2]) and fundamental frequency (F0) across prosodic domain-initial positions to examine whether these phonetic properties are also affected by the different prosodic positions.

The main purpose of my dissertation was to test if the speech clarity according to prosodic position and word type affects the listener's perception of the segment in real speech. With perception study, I examined whether the acoustic properties in domain-initial segments are perceptible prosodic markers that can be perceptually distinguished as a function of different prosodic position and word type. In order to investigate whether the acoustic properties of the prosodic domain-initial positions affect the perception of the target segments, the perceptual differences of CVs across different prosodic levels were compared by manipulating the prosodically conditioned properties of target CVs.

My claim is that speakers focus their effort on clearly articulating domain-initial segments to make them more highly perceptible to listeners, and that listeners use these acoustic correlates of the domain-initial segments in auditory word recognition. The enhanced accessibility of segmental information in prosodic domain-initial positions would be helpful for the recognition of word-initial segments at domain beginnings where less top-down (e.g., syntactic and semantic) information is available. In addition, the enhanced phonetic properties of initial segments in the prosodic domains could possibly tell the listener about the strength of the prosodic boundary.

1.5 Outline of the Study

The chapters that follow present production and perception experiments designed to examine prosodic domain-initial properties of Korean stops and fricatives. In Chapter 2, the phonetic properties of stops and fricatives are compared in IP, PP and Wd-initial positions. I also compare the acoustic properties of the prosodic domain-initial segments in real words to those in nonsense words. The differences are tested in terms of durational parameters and spectral parameters. The durational parameters include VOT for stops, fricative duration for fricatives and CV duration for both stops and fricatives. In addition to the prosodic domain-initial properties, the vowel duration of prosodic domain-final position is also tested. The spectral parameters include relative burst energy for stops and centroid frequency for fricatives. F0 and [H1-H2] are also tested for both stops and fricatives. The production study in Chapter 2 serves as the basis for the perception study in Chapter 3.

Chapter 3 tests the perception of a manipulated CV syllable from three different prosodic domain-initial positions. The prosodic properties of the target syllable are altered to find out whether listeners are sensitive to the phonetic properties driven by the different prosodic domain-initial positions. The identification of mismatched CVs is also compared between real and nonsense words. Chapter 4 summarizes the main findings and arguments of each chapter. Moreover, the chapter provides some suggestions for future research.

Chapter 2: The Acoustic Properties of Korean Stops and Fricatives in Prosodic Domain-initial Positions

2.1 Objectives

The goal of this production study was to confirm whether initial stops and fricatives in higher prosodic domains have more enhanced acoustic properties than those in lower prosodic domains. I examined how prosodic domains are acoustically distinguished among IP, PP and Wd levels in Korean stops and fricatives.

In addition, I examined whether the acoustic properties of prosodic domain-initial segments are affected by word types such as real and nonsense words. I compared the acoustic properties of real words to those of nonsense words in each prosodic domain-initial position.

In order to compare the acoustic properties across prosodic positions and between word types, durational and spectral parameters for Korean stops and fricatives were considered. The durational parameters included VOT for stops, fricative duration for fricatives, and CV duration for both stops and fricatives. The spectral parameters consisted of relative RMS burst energy for stops, centroid frequency for fricatives, and fundamental frequency (F0) and [H1-H2] difference for both stops and fricatives. In addition to the prosodic domain-initial properties, final vowel duration was also compared across prosodic domains.

Before presenting the specific hypotheses in this study, I briefly sum up the results of acoustic analyses of prosodic domain-initial effects on Korean stops and fricatives. Korean stops in higher prosodic domain-initial positions were shown to be progressively lengthened relative to consonants in lower prosodic domain-initial positions (i.e., VOT, total voiceless interval, closure duration) as a function of prosodic hierarchy (Cho & Jun (2000), Cho & Keating (2001)). The results of RMS burst energy for Korean stops were shown to be smaller for Ui and IPi than APi and Wi for /t^h, t'/ in Cho & Keating (2001) although they were not statistically significant. The RMS burst energy of /t/ in English was also found to be lower in Utterance-initial position than in Utterance-medial position due to low level of intraoral pressure.

In the acoustic analyses of prosodic domain effects on Korean fricatives (S. Kim (2001), fricative duration was shown to be longer in higher prosodic domains than in lower prosodic domains. The centroid frequency for fricatives was found to be higher in higher prosodic domains than in lower prosodic domains.

The prosodic domain-initial effect on the following vowel of stops and fricatives has not been studied but F0 of the following vowel is expected to be higher in higher prosodic domains than in lower prosodic domains due to the effect of declination of F0 in utterance. [H1-H2], on the other hand, demonstrated the opposite trend: S. Kim (2001) showed that [H1-H2] following fortis fricatives was lower in higher prosodic domain-initial positions than in lower prosodic domain-initial positions. In my pilot study, [H1-H2] values after stops were also found to be greater in lower prosodic domain-initial positions than in higher prosodic domain-initial positions, although between-subject variations exist in the results.

With regard to the properties of the right edge of prosodic domains, phrase final lengthening has been shown in several studies and regarded as the realization of prosodic boundary marker in addition to boundary tone. Final vowel was lengthened cumulatively when the prosodic domain gets higher (Cho & Keating (2001)).

Contrary to the prosodic domain-initial effect, there were few studies comparing the phonetic properties of segments across prosodic domains between real and nonsense words. But, the above acoustic properties are expected to be more reinforced in nonsense words than in real words due to the hyperarticulation of target words produced as nonsense words.

Based on the results from the previous studies of prosodic domain-initial effects on Korean stops and fricatives, I propose the following research hypotheses for this experiment as in (1).

(1) Hypotheses

(i) Stops

(a) VOT is greater in higher prosodic domain-initial positions than in lower prosodic domain-initial positions.

(b) CV duration is greater in higher prosodic domain-initial positions than in lower prosodic domain-initial positions.

(c) Relative burst energy is lower in higher prosodic domains than in lower prosodic domain-initial positions.

(d) Fundamental frequency of vowels after target stops is higher in higher prosodic domain-initial positions than in lower prosodic domain-initial positions.

(e) [H1-H2] values of vowels after target stops are lower in higher prosodic domain-initial positions than in lower prosodic domain-initial positions.

(f) The duration of a vowel before a boundary in the higher prosodic domains is longer than that in the lower prosodic domains.

(ii) Fricatives

(a) Fricative duration is greater in higher prosodic domain-initial positions than in lower prosodic domain-initial positions.

(b) CV duration is greater in higher prosodic domain-initial positions than in lower prosodic domain-initial positions.

(c) Centroid frequency is higher in higher prosodic domain-initial positions than in lower prosodic domain-initial positions.

(d) Fundamental frequency of vowels after target fricatives is higher in higher domain-initial positions than in lower domain-initial positions.

(e) [H1-H2] values of vowels after target fricatives are lower in higher prosodic domain-initial positions than in lower prosodic domain-initial positions.

(f) The duration of the vowel before a boundary in the higher prosodic domains is longer than that in the lower prosodic domains.

(iii) The prosodic domain-initial segments from nonsense words have more distinctive acoustic properties than those from real words.

2.2 Methods

2.2.1 Participants

Three male and three female speakers, aged 30 to 43, participated in the study. They are speakers of Seoul Korean, with no known hearing problems. They were graduate and undergraduate students at the University of Texas at Austin.

2.2.2 Materials

In order to test the hypotheses, the stimuli consisted of real words and nonsense words in which each of the three stops /t, t^h, t'/, and each of two fricatives /s, s'/, appeared in IP, PP, and Wd-initial positions. To place the target segments in IP-initial position, the target segments were placed after vocative words followed by a comma. It has been noted that parentheticals, non-restrictive relative clauses, preposed adverbials, tag questions, expletives, vocatives, and certain moved elements are produced with their own Intonational Phrases (Selkirk (1986), Nespor & Vogel (1986)). Since a vocative word formed its own IP, the following segments were placed in the initial position of following IP as in table 2.1(a). The IP was followed by a pause and also marked by lengthening on the final vowel and boundary tone. To place the target segments in PP-initial position, the segments were placed in the initial position of a predicate phrase as in table 2.1(b). The predicate phrase made its own PP. The preceding subject noun phrase was marked by either LHLH or HHLH[16] phrasal tone but since the predicate phrase was in an IP-final position, it was marked by boundary tone (L%). To place the target segments in Wd-initial position, I placed the segments in the initial position of a word in the middle of an object phrase. The object NP formed a PP or made a PP with a following verb phrase. When the object NP formed a PP, it was marked by phrasal tone (LH), but

[16] The PPs start with H when the target segments are tense, aspirated stops, or fricatives. It has already been noted in Jun (2000).

when it made a PP with a following verb phrase, it ended with a IP boundary tone L%. So, the target segments were placed in the middle of the PP.

In order to control vowel context, the low vowel /a/ was used as the preceding and following vowel of the initial consonant of each prosodic domain. All test sentences consisted of 13 syllables in order to control the speech length. To summarize, there were 30 conditions (5 segment types * 3 prosodic positions * 2 word types) and 3 test sentences are used for each condition. Filler utterances, /ʧ, $ʧ^h$, ʧ'[17]/ were used in the same prosodic positions both in real and nonsense words (3 segment types * 3 prosodic positions * 2 word types * 2 test sentences). There were two repetitions of those sentences, yielding a total of 252 sentences per speaker. Representative stimuli are shown in table 2.1. The target CV is underlined in each prosodic position below.

[17] /ʧ'/ represents tense voiceless affricate in Korean.

Table 2.1 Example stimuli

a. IP-initial position:					
[IP nae doŋseŋdɯ-ɾa],		[IP [PP tasugjəllo		kjəltʃ’əŋ-he]]	
my brothers-Voc		decision by majority		decide-IMP	
‘My brothers, decide by majority’					
b. PP-initial position:					
[IP [PP sosimhan	minsu-ka]	[PP tasi	todʒən-ɯl	he-s’ə]]	
Timid	Minsu-Nom	again	challenge-Acc	do-past ending	
‘A timid Minsu challenged again’					
c. Wd-initial position					
[IP [PP minjəŋi-ka]	[PP [Wd op’a]	[Wd tat^hɯ-lɯl]	pəɾjə-s’ə-yo]]		
Minyoung-Nom	brother (‘s)	dart-Acc	throw –past-DEC		
‘Minyoung threw her brother’s dart’					

In order to ensure that any enhanced phonetic properties of the target segments are due to the effect of prosodic domain-initial position, other prosodic factors need to be controlled. For example, Hay et al. (2006) noted that there are several distinctiveness-enhancing correlates of vowels in [+focus] context in English, French and German. In producing vowels in [+focus] context, all three language groups increased spectral differences among vowels and German speakers increased vowel duration differences. So, to avoid possible confounding effects from focus, wh-questions and focus-cueing sentences were given in parenthesis before the target sentence. Since subjects tend to put their contrastive focus on the subject phrase as an answer to a wh-word, the target words were controlled in non-focused position. To avoid focus on the target IP, a focus-cueing sentence was suggested before the target sentence. The focus-cueing sentence was exactly the same as the target sentence except for the words in the first IP. Because the target IP already had the same information, speakers tended to put a contrastive focus on the initial IP with new information and the target word in the second IP was thus controlled in non-focused position.

With respect to the focus effect in the phonological phrasing, it was noted that a focused word starts a new AP (or PP) and includes all following words within an Intonational Phrase, unless one of those following words itself is focused (Jun (1993), Jun & Lee (1998)). However, Jun & Lee (1998) suggested that not all speakers produced the focused word and the following words as one phrase and that the dephrasing occurred more frequently as the phrase became shorter. In my pilot study of focus effect on phonological phrasing in Seoul Korean, I found that most of the utterances did not show dephrasing after focused words. In the current study, the dephrasing after focused words was shown in relatively very small number of utterances and I excluded the utterances when they had different phonological phrasing that changed the place of target segments in the prosodic domains.

In the recording, subjects were not asked to read the wh-questions and focus-cueing sentences. They were asked to read the test sentences as in (2a') and (2b'). The sample focus-cueing utterances and wh-questions are shown in (2a) and (2b), respectively.

(2) a. Nae ʧʰingudɯ-ɾa t'asɯhage ibə-ɾa

My friends-VOC, warm take on-IMP

'My friends, take your warm clothes'

a' Nae donseŋdɯ-ɾa t'asɯhage ibə-ɾa

My brothers, warm take on-IMP

'My brothers, take your warm clothes'

b. nu-ka tasi toʤən-ɯl he-s'ə

Who-Nom again challenge-Acc do-interrogative marker

'Who challenged again?'

b'. sosimhan minsu-ka <u>ta</u>si toʤən-ɯl he-s'ə

Timid Minsu-Nom again challenge-Acc do-past ending

'A timid Minsu challenged again'

2.2.3 Procedures

Subjects were asked to read materials written in Korean orthography, at a self selected speaking rate throughout the recording session. The test sentences were presented in random order on separate slides in a timed PowerPoint presentation on an IBM laptop. The word list was rehearsed with sample sentences before the recording. Speakers were recorded in a sound-proof booth, using a solid state recorder, Maranz PMD 670, in the Phonetics Laboratory at the University of Texas at Austin. Recorded materials were digitized at a sampling rate of 22050 Hz.

2.2.4 Measurements

In order to compare the acoustic properties of stops and fricatives in IP, PP and Wd-initial positions, I measured VOT and relative RMS burst energy for stops, and fricative duration and centroid frequency for fricatives. In addition, I measured CV duration, final vowel duration, F0 and [H1-H2] for both stops and fricatives. All measurements were taken in Praat. The detailed measurement points are specified in (3) and (4).

(3) Stops

(i) Durational Parameters

a. Voice onset time:

VOT for /t, t^h, t'/ was taken from the point of release of noise or aperiodic wave to onset of periodicity in waveform

b. CV duration:

CV duration for stops was taken from the point of the stop release to the offset of the vocalic energy associated with the following vowel. For stops, the duration of the silent stop closure was not included in measuring CV duration because in IP-initial position, the entire silent period before the stop release is likely to include a silent period associated with the IP boundary that is independent of stop production. So, the differences in stop closure duration among each prosodic position were eliminated in this study.

c. Final vowel duration:

The vowel duration was measured from the point in the expanded waveform at which waveform amplitude and complexity begin to rise to the point where the decline in waveform amplitude and complexity end.

(ii) Spectral Parameters

d. RMS burst energy:

The acoustic energy at the burst was measured from an FFT spectrum giving the RMS value over all frequencies. A 10 ms window was centered over the release of the stop /t, t^h/. For the fortis stop /t'/, a shorter window (less than 5 ms) was used in order to prevent the window from including the following vocalic energy. To normalize the energy difference across speakers, the percentage value of the burst energy relative to the energy at the midpoint of the following vowel was measured.

e. Fundamental frequency (F0):

F0 was taken at the midpoint of the following vowel, using the pitch tracking function in Praat. When the pitch line abruptly moved or was discontinued, F0 was calculated by measuring the duration of the relevant period in seconds. As supplementary checks, the tenth harmonic values were divided by 10 from an FFT with a 25 ms window.

f. [H1-H2]:

The amplitude (dB) difference between the first (H1) and the second (H2) harmonic was measured just after the first full glottal pulse of the vowel onset in the waveform. The amplitude values were calculated using a narrowband fast Fourier transform spectrum using a Hamming window (window length of 25 ms).

(4) Fricatives

(i) Durational Parameters

a. Fricative duration:

The fricative duration was taken from the beginning of high frequency noise to the onset of periodicity in the following vowel, using both the spectrogram and waveform.

b. CV duration:

CV duration for /s, s'/ were taken from the point of the beginning of the high frequency frication noise to the offset of the vocalic energy associated with the following vowel.

c. Final vowel duration:

The vowel duration was measured from the point in the expanded waveform at which waveform amplitude and complexity begin to rise to the point where the decline in waveform amplitude and complexity end.

(ii) Spectral Parameters

d. Centroid of the fricative noise:

Centroid values were taken from FFT spectra using a 25 ms window centered around the midpoint of the fricative portion.

e. F0:

F0 was taken at the midpoint of the following vowel, using the pitch tracking function in PRAAT. When the pitch line abruptly moved or was discontinued, F0 was calculated by measuring the duration of the relevant period in seconds. As supplementary checks, the tenth harmonic values were divided by 10 from an FFT with a 25 ms window.

f. [H1-H2]:

The amplitude (dB) difference between the first (H1) and the second (H2) harmonic was measured just after the first full glottal pulse of the vowel onset in the waveform. The amplitude values were calculated using a narrowband fast Fourier transform spectrum using a Hamming window (window length of 25 ms).

2.2.5 Statistical Design

A repeated measures analysis of variance (ANOVA) was performed with three within-subjects factors (segment, prosodic position, and word type) using SPSS 1.6. The prosodic position factor consisted of three different levels: IP, PP and Wd-initial positions. The word type factor was composed of two different levels: real and nonsense words. For stops, the dependent variables were VOT, CV duration, final vowel duration, relative RMS burst energy, F0 and [H1-H2]. For fricatives, the dependent variables were fricative duration, CV duration, final vowel duration, centroid frequency, F0 and [H1-H2].

2.3 Results

2.3.1 Stops

2.3.1.1 Durational Parameters

In the results of duration parameters, I excluded several test sentences for /t/ in Wd-initial position due to the voicing of the segment. It has been known that lenis stops in Korean become voiced intervocalically within an intermediate phrase, AP (Jun (1993) or PP (Kang (1992)). However, only one speaker produced a voiced lenis stop in Wd-initial positions in this experiment.

2.3.1.1.1 Voice Onset Time

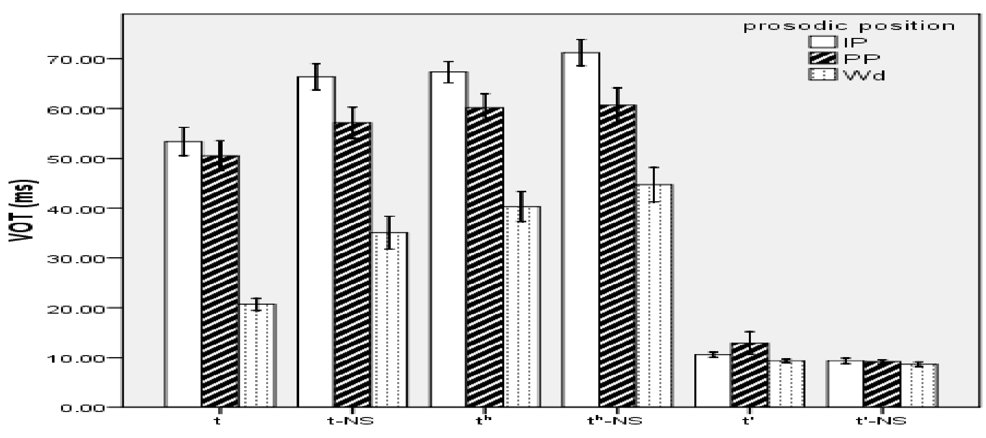


Figure 2.1 Pooled graph for VOT by segment type (combined with word type) * prosodic position

In Figure 2.1, VOT values of the stop categories in real and nonsense words (ms) are compared across three different prosodic domain-initial positions IP, PP and Wd. In the graph, t-NS, t^h-NS, and t'-NS represent /t/, /t^h/ and /t'/ in nonsense words, respectively. It displays that for /t/ and /t^h/, VOT values are greatest in IP, intermediate in PP and smallest in Wd-initial position. For /t'/, there is no significant distinction among the three different domain-initial positions and between two word types.

The graphs from individual speakers are presented in Appendix A.1.(1). All speakers show the same pattern as in the pooled graph except for the speakers F3 and M2. The speaker F3 produced longer VOT in PP-initial position than in IP initial position for /t/ and /t^h/ in the real word condition, and both F3 and M2 exhibited no difference between IP and PP for /t^h/ in the nonsense word condition.

Results of a repeated measures analysis of variance (ANOVA) on VOT showed significant main effects for stop category ($F (2, 10) = 55.581$, $p < 0.000$), prosodic position ($F (2, 10) = 8.989$, $p = .006$), and word type ($F (1, 5) =$, $p = .009$). There was also a highly significant interaction among prosodic position, stop category and word type ($F (4, 20) = 32.546$, $p < .000$).

Pairwise post hoc comparisons in Table 2.2 revealed that VOT values for both /t/ and /t^h/ between IP and Wd, and PP and Wd were significantly different in real and nonsense words. However, there was no significant difference of VOT between IP and PP for /t^h/ in real and nonsense word, and for /t/ in real words. /t/ in nonsense words showed a significant difference between IP and PP. That is, the VOT distinction between IP and PP was not as clear as between IP and Wd, and between PP and Wd. The results showed the pattern IP, PP ＞＞ Wd in general.

In the comparison between real and nonsense words, only the lenis stop showed a significant difference between the two word types in IP–initial position (p=.008). But there were no significant differences between the two word types in the other prosodic positions for /t/ and /t^h/. So, speakers exhibited enhanced VOT in higher prosodic domain-initial positions for /t/ and /t^h/ and a limited word type effect on /t/. That is, the effect of prosodic position interacts with the effects of stop category and word type.

Table 2.2 The results of post hoc tests for VOT

	Real Word		Nonsense	
/t/	IP, PP	p = .413	IP, PP	p= .022
	IP, Wd	p <.000	IP, Wd	p= .001
	PP, Wd	p < .000	PP, Wd	p= .007
/t^h/	IP, PP	p = .124	IP, PP	p = .076
	IP, Wd	p= .001	IP, Wd	p= .003
	PP, Wd	p < .000	PP, Wd	p= .022

2.3.1.1.2 CV Duration

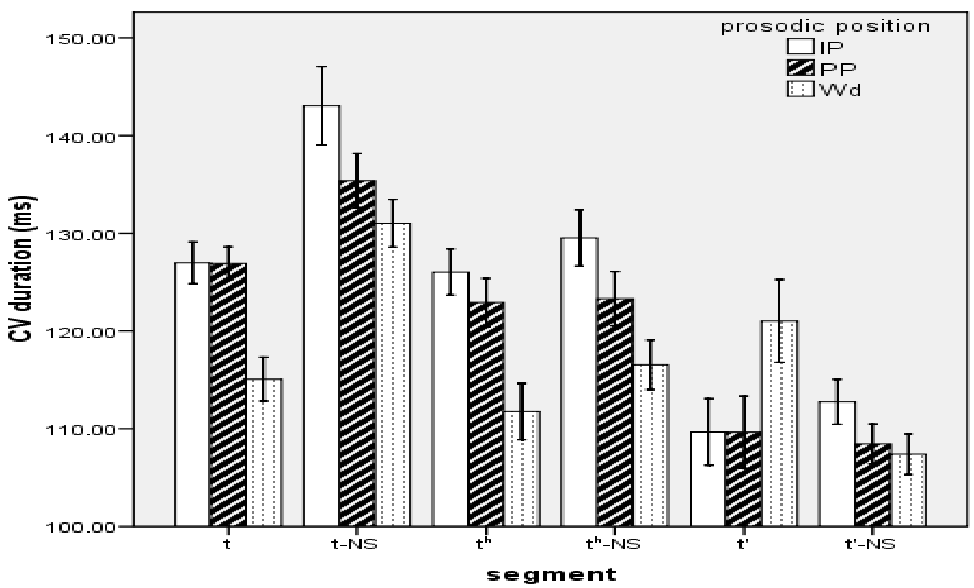


Figure 2.2 Pooled graph for CV duration by segment type (combined with word type) * prosodic position

In Fig.2.2, the CV durations of target stops were compared at IP, PP and Wd-initial positions between the two word types. The pooled graph shows that the CV duration is shortest in Wd-initial position than in the other prosodic domain-initial positions for /t/ and /t^h/ in real and nonsense words, and /t'/ in nonsense words. These segments show cumulatively increasing CV durations when a position moves up in prosodic hierarchy. But there was no difference of CV durations between IP and PP-initial positions for /t/ in real words. Relative to the other stops, /t'/ in real words showed a different pattern in CV duration. It had longer CV duration in Wd-initial position than in the other higher prosodic domain-initial positions and the CV durations of /t'/ in real and nonsense words were relatively shorter than those of the other stops. The duration of the stop closure was known to be longest for fortis stops among Korean stop categories, but the acoustic closure duration was excluded in measurements of CV duration, resulting in a relatively shorter CV duration for fortis stops. So, the inconsistent variation for fortis stops across prosodic domains is related to the results of VOT.

However, the graphs of the individual speakers in Appendix A.1.(2) do not show the same pattern as the pooled graph. Only M2 shows the same pattern as the pooled graph but for the rest speakers, CV duration in higher prosodic domain-initial positions was not consistently longer than in lower prosodic domain-initial positions for all stop categories in real and nonsense words. Relative to the results of VOT, the results of CV duration showed a less consistent pattern across speakers.

The main effects for stop category and word type were significant but the prosodic position effect was not significant ($F(2, 10) = 25.864$, $p < .000$ for stop category; $F(2, 10) = .205$, $p = .818$ for prosodic position; $F(1, 5) = 7.980$, $p = .037$ for word type). But there was highly significant interaction of prosodic position, stop category and word type ($F(4, 20) = 7.147$, $p < .001$).

The results of post hoc tests in Table 2.3 showed that for /t/ and /t^h/, the CV durations across the three prosodic positions were significantly different only in the real word condition ($p=.006$ for /t/ and $p=.001$ for /t^h/). In this context, CV durations differed significantly between IP and Wd, and between PP and Wd-initial positions. But they did not differ significantly between IP and PP-initial positions for both /t/ and /t^h/. Likewise the results of VOT, the results of CV duration for lenis and aspirated stops showed the pattern IP, PP >> Wd in real words. However, there was no significant difference across the three prosodic positions for /t'/ in real ($p=.105$) and nonsense words ($p=.115$).

Table 2.3 The results of post hoc tests for CV duration

Real Word	/t/	/tʰ/
IP, PP	p=.934	p=.380
IP, Wd	p=.019	p=.005
PP, Wd	p=.026	p=.003

The post hoc tests for the two word types revealed that only lenis stop showed significant difference between real and nonsense words in Wd-initial position (p=.007). The other stops did not show significant difference between the two word types across different prosodic positions. Thus, the effect of prosodic position on CV duration depends on the effect of stop category and word type.

2.3.1.1.3 Final Vowel Duration

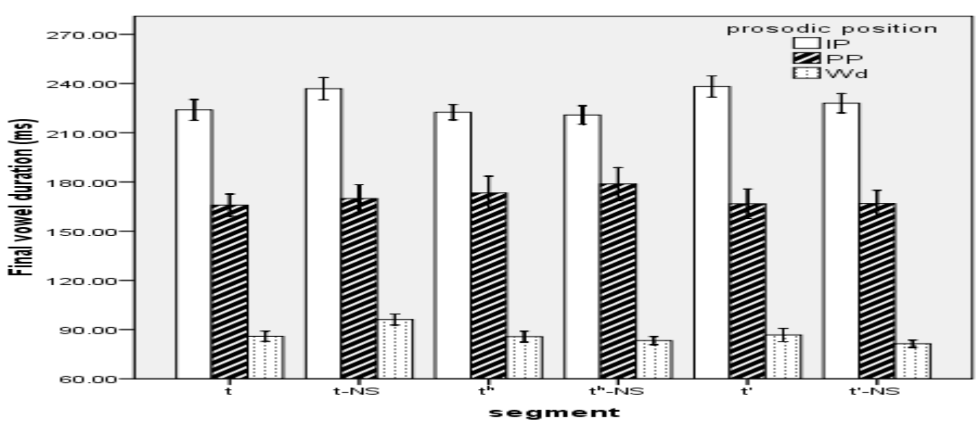


Figure 2.3 Pooled graph for final vowel duration by segment type (combined with word type) * prosodic position

In Fig. 2.3, the vowel durations are compared at each prosodic domain-final position. It is seen that for all stop categories, the final vowels are the longest in IP, intermediate in PP and shortest in Wd. The graphs from individual speakers presented in Appendix A.1.(3) show that speaker M3 has longer vowel duration in PP-final position than in the other prosodic domain-final positions for aspirated stops. For M2, the difference between IP and PP is relatively small for all stops.

The main effect for prosodic position was significant but the other main effects were not significant ($F(2, 10) = .426$, $p = .664$ for stop category; $F(2, 10) = 30.748$, $p < .000$ for prosodic position; $F(1, 5) = .060$, $p = .816$ for word type). The interaction of three factors was not significant ($p=.933$). But, the interaction of segment and word type, and that of prosodic position and word type were significant ($p=.017$ for stop category * word type, $p < 0.000$ for prosodic position * word type). Different from the results in VOT and CV durations, the final vowel duration did not show difference across phonation types. The significant interaction was caused by the highly significant effect of prosodic position in final vowel duration.

The Pairwise post hoc tests in Table 2.4 displays that all prosodic domains are distinguished by the final vowel duration except for /t^h/ between IP and PP in real and nonsense words. The final vowel durations show the pattern IP >> PP>> Wd before /t/ and /t'/, while before /t^h/, they show the pattern IP, PP >> Wd. Due to longer vowel duration in PP-final position from the speaker M3, the distinction between IP and PP in the pooled result is rendered less distinctive.

Table 2.4 The results of post hoc tests for final vowel duration

	Real Word	Nonsense Word
/t/	IP, PP : p=.043 IP, Wd: p<.000 PP,Wd: p=.006	IP, PP: p=.013 IP, Wd: p <.000 PP,Wd: p=.018
/t^h/	IP, PP: p=.113 IP, Wd: p<.000 PP,Wd: p=.012	IP, PP: p=.141 IP, Wd: p<.000 PP,Wd: p=.010
/t'/	IP, PP: p=.023 IP, Wd: p<.000 PP,Wd: p=.019	IP, PP: p=.039 IP, Wd: p<.000 PP,Wd: p=.009

2.3.1.2 Spectral Parameters

2.3.1.2.1 Relative RMS Burst Energy

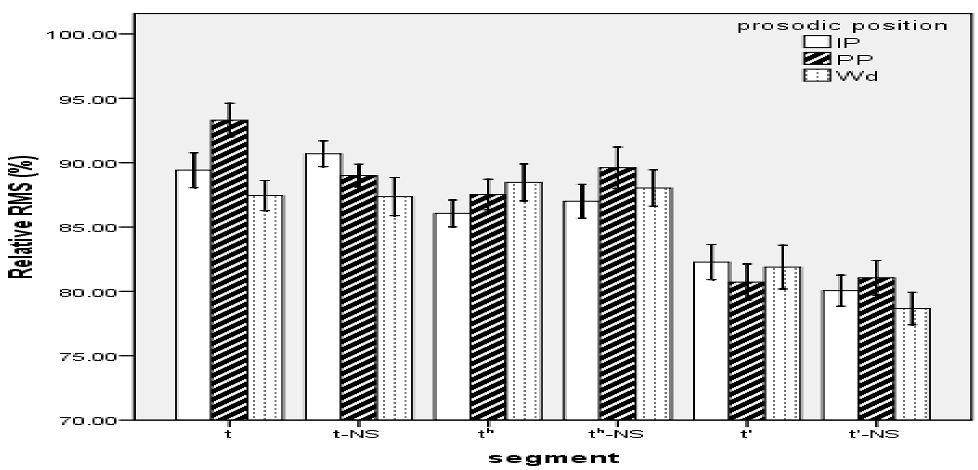


Figure 2.4 Pooled graph for relative RMS burst energy by segment type (combined with word type) * prosodic position

Figure 2.4 displays that the results of relative RMS burst energy do not show systematic difference across the three prosodic positions for all stops but that /t'/ has relatively smaller relative RMS burst energy than the other stops. The individual graphs from all speakers in Appendix A.1.(4) also reveal that each speaker shows a different pattern in the results of relative RMS burst energy.

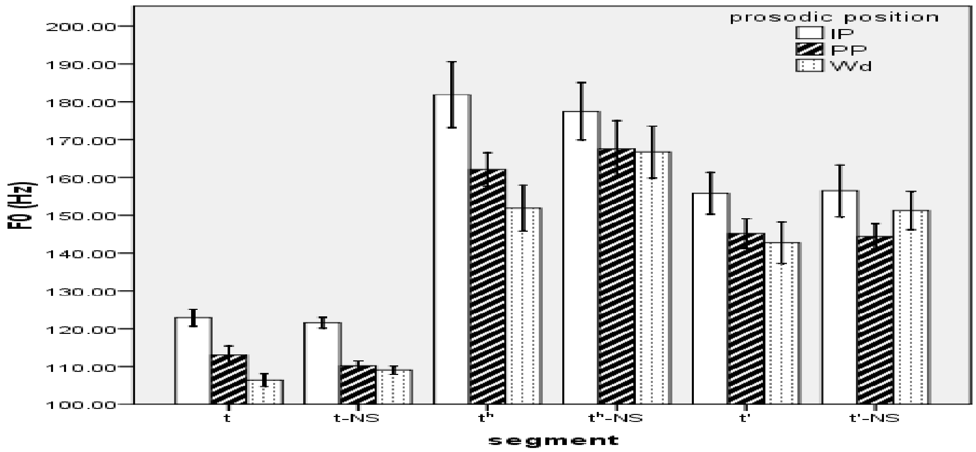
The main effect for stop category was significant but the other main effects were not significant ($F(2, 10) = 27.738$, $p < .000$ for stop category; $F(2, 10) = .628$, $p = .553$ for prosodic position; $F(1, 5) = .021$, $p = .890$ for word type). None of the interactions were significant. The significant effect of stop category was due to the fact that that the relative RMS of /t'/ was relatively smaller than that of the other stops.

Pairwise post hoc comparisons revealed a significant difference between fortis and lenis stops ($p=.001$) and fortis and aspirated stops ($p< .000$). There was no significant difference between lenis and aspirated stops ($p =.211$).

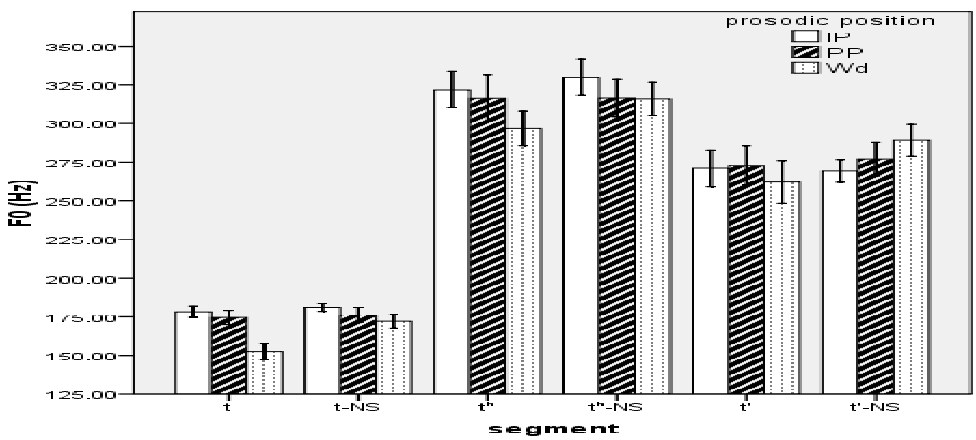
The results of this study did not have the same pattern as in Cho & Keating (2001). They showed that RMS burst energy was smaller for Ui and IPi than for APi and Wi for /t^h, t'/ but there was no systematic pattern for /t/. Only /t^h/ in the current study paralleled their findings, but /t'/ did not replicate their results in Cho & Keating (2001). As noted in chapter 1, the results were not statistically significant. The significant difference across prosodic positions was also not found in this study with a relatively large subject pool. So, this result does not support the hypothesis in that the relative RMS burst energy is lower in higher prosodic domain-initial positions than in lower prosodic domain-initial positions.

With regard to the segmental difference of RMS burst energy, aspirated stops were reported to have significantly greater burst energy than the other stop categories but there was no significant difference between the fortis and the lenis stops (Cho et al. (2002)). However, the findings of this study were not in accord with the results of the previous studies. The relative RMS burst energy was greater for lenis and aspirated stops than for fortis stops and there was no consistent difference between lenis and aspirated stops. This pattern seems to be caused that aspirated and lenis stops are produced with greater intraoral airflow than fortis stops (Cho et al. (2002)). Since the target stops in this study were placed in the middle of utterance, the relative RMS burst energy was relatively high for all stops because of greater intraoral airflow in the middle of utterance. But this study does not fully explain the greater variation of relative RMS burst energy for lenis stops.

2.3.1.2.2 Fundamental Frequency



a. Pooled graph for male speakers



b. Pooled graph for female speakers

Figure 2.5 Pooled graph for F0 by segment type (combined with word type) * prosodic position

The pooled graphs from male and female speakers are shown in Figure 2.5. F0 after a lenis stop is lower than F0 after the other stops, in general. F0 in IP-initial position is higher than in the other prosodic domain-initial positions for all stops in male speakers and for lenis and aspirated stops in female speakers. The graphs from individual speakers in Appendix A.1.(5) display that the speaker M1, M2 and F2 show higher F0 in IP-initial position than in the other lower prosodic domain-initial positions but there is variation in F0 across PP and Wd-initial positions. The rest speakers showed less consistent variation across prosodic domain-initial positions.

The main effect for stop category was significant but the effect of prosodic position was not significant ($F(2, 10) = 13.767$, $p = .001$ for stop category; $F(2, 10) = 1.444$, $p = .281$ for prosodic position). But the main effect for word type was significant ($F(1, 5) = 12.391$, $p = .017$ for word type). None of the interactions were significant.

Pairwise post hoc comparisons showed that f0 after lenis stops was significantly lower than after aspirated stops ($p=.012$) and fortis stops ($p=.019$). The difference of F0 between aspirated stops and fortis stops was significant ($p=0.13$). Although there was variation across prosodic domains, F0 in nonsense words was higher than that in real words.

In comparison of F0 after stop categories, there was no distinction of F0 between aspirated and fortis stops for Seoul speakers in Cho et al. (2002). But in this study, the target stops were placed in the initial position of a word in isolation and words produced in isolation tend to have greater duration and amplitude than words in the middle of utterance. The hyperarticulation of target segments in isolated words might cause less distinction of F0 between aspirated and fortis stops. However, target stops in the current study were placed across prosodic domain-initial positions in the middle of test sentences. In addition, target segments were controlled to be placed in non-focused positions. Since all speakers in this study produced greater F0 after aspirated stops than after fortis stops, the results of this study show that the distinction of F0 after aspirated and fortis stops exists in continuous speech.

2.3.1.2.3 [H1-H2]

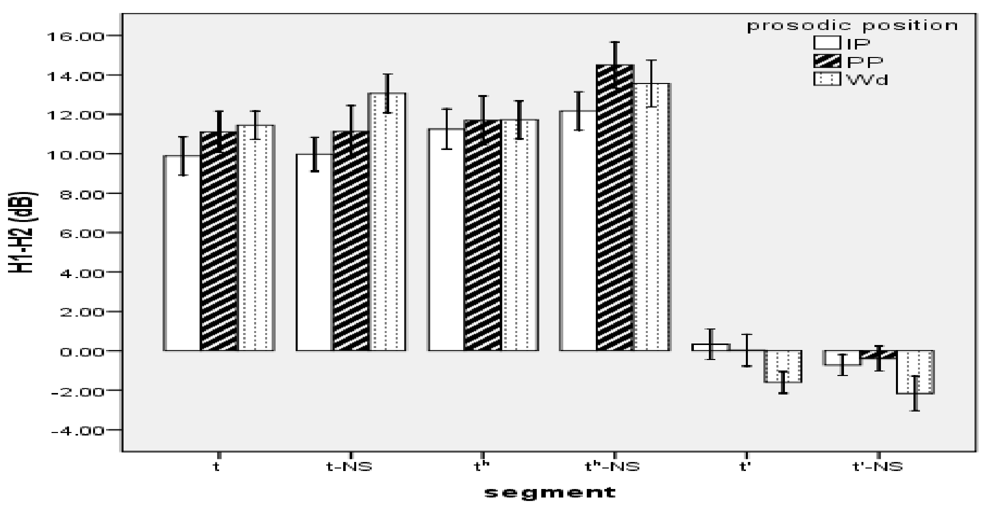


Figure 2.6 Pooled graph for [H1-H2] by segment type (combined with word type) * prosodic position

In Figure 2.6, [H1-H2] values at the following vowel of target stops are compared at IP, PP and Wd-initial positions in both real and nonsense words. The pooled graph illustrates that [H1-H2] values are greater (positive) after lenis and aspirated stops than after fortis stops. For /t/ and /t^h/, [H1-H2] values in IP-initial position are smaller than those in the lower prosodic positions. For /t'/, [H1-H2] values are smaller and more negative in Wd-initial position than in the higher prosodic positions. However, the individual speakers do not show the same pattern, except for speaker F2 as shown in Appendix A.1.(6).

The main effect for stop category was significant but the other main effects were not significant ($F(2, 10) = 17.450$, $p = .001$ for stop category; $F(2, 10) = 1.479$, $p = .274$ for prosodic position; $F(1, 5) = .009$, $p = .929$ for word type). None of the interactions were significant. The significant effect of stop category was caused by the difference of [H1-H2] between fortis stops, and aspirated and lenis stops. There was no significant difference in [H1-H2] between lenis and aspirated stops ($p=.619$) but [H1-H2] after fortis stops was significantly different from lenis stops ($p=.013$) and aspirated stops ($p <.000$). To sum up, the stop categories were differentiated in part by the voice quality of the following vowel but there was no effect of prosodic position in the vowel quality of the following vowel.

With regard to the distinction across stop category, Cho et.al (2002) reported that all Seoul speakers showed a pattern of fortis < aspirated < lenis in the results of [H1-H2].

Ahn (1999) showed an order of fortis < lenis < aspirated stops in both raw and normalized [H1-H2] values. Cho et. al (2002) mentioned that the difference could be due to the relatively younger speakers (aged from 29 to 43) in Ahn's study. However, the speakers in the current study aged from 30 to 43 and showed an order of fortis < lenis, aspirated stops in the results of [H1-H2]. I think there is no systematic distinction in voice quality of the following vowel between lenis and aspirated stops. It can be assumed that the voice quality only provides a clue to distinguish fortis stops from lenis and aspirated stops in Korean.

2.3.2 Fricatives

2.3.2.1 Durational Parameters

In the results of duration parameters, I excluded several test utterances for /s/ in Wd-initial position due to the voicing of the segment. Cho et al. (2002) already noted that /s/ in Word-initial position showed voicing between vowels although voicing of /s/ occured in a gradient fashion. In this experiment, only one speaker showed intervocalic voicing of /s/ in half of the tokens.

2.3.2.1.1 Fricative Duration

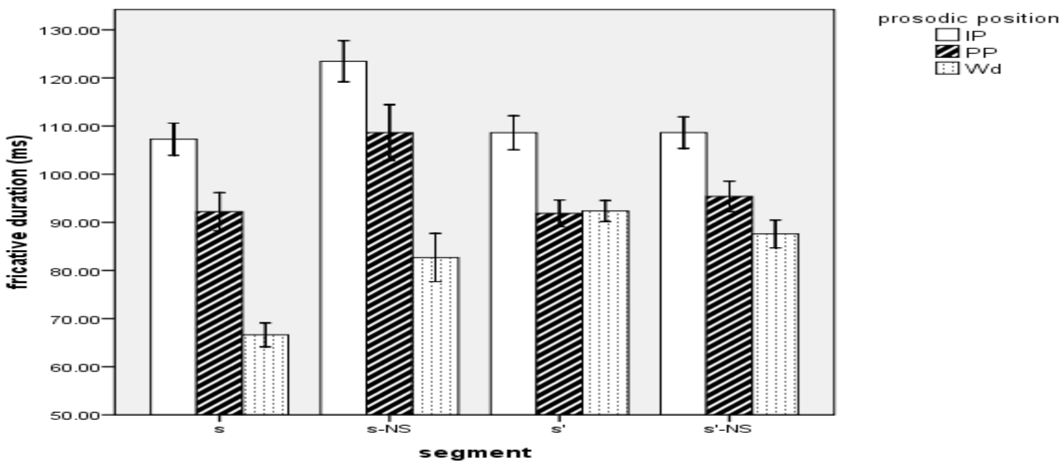


Figure 2.7 Pooled graph for fricative duration by segment type (combined with word type) * prosodic position

Figure 2.7 illustrates fricative durations of fricative categories in real and nonsense words across IP, PP and Wd-initial positions. The fricative duration is longest in IP-initial position, intermediate in PP-initial position and shortest in Wd-initial position, except for /s'/ in real words. /s'/ in real words does not show a difference between PP and Wd-initial position. The individual graphs from all speakers are presented in Appendix A.2.(1). For lenis fricatives, most speakers showed the same pattern as in pooled graph except for speaker F3. The speaker F3 has a longer fricative duration in PP-initial position than in IP and Wd-initial positions similar to the results for

VOT. For fortis fricatives, most speakers did not show consistent variation across prosodic domains except for M2.

There was no main effect of fricative category ($F(1, 5) = .053$, $p < .827$), prosodic position ($F(2, 10) = 2.166$, $p = .165$) and word type ($F(1, 5) = .474$, $p = .522$). But, the interaction of prosodic position, fricative category and word type was highly significant ($F(1, 10) = 20.295$, $p < .000$).

The results of pairwise post hoc analyses are seen in Table 2.5. For /s/ in real and nonsense words, there was no significant difference between IP and PP-initial positions but there was a significant difference between IP and Wd, and PP and Wd-initial positions. For fortis fricatives, there was a significant difference between IP and PP in both real and nonsense words, and IP and Wd in nonsense words. But /s'/ did not show difference between IP and Wd levels in real words, and between PP and Wd levels in real and nonsense words.

The comparison between real and nonsense words revealed that only lenis fricatives exhibited significant difference between the two word types in IP-initial position ($p=.019$). In short, the effect of prosodic position interacts with the effects of fricative category and word type since fortis fricatives showed less variability across prosodic positions and between the two word types relative to lenis fricatives.

The results of lenis fricatives in S. Kim (2001) also presented increasing fricative duration when the prosodic domain gets higher but for fortis fricatives, only one of the two speakers showed progressively increasing fricative duration in higher prosodic domain-initial positions. The between speaker variation seems to be caused by variation between phonation types. As seen in the durational variation of stops, lenis and aspirated categories revealed enhanced duration as a function of prosodic position, while fortis stops did not show any durational variation. Likewise, for fortis fricatives, the less consistent variation across prosodic position seems to be caused by the properties of fortis category. Alternatively, since S. Kim (2001) did not control focus on the target segments, the between subjects variance might be caused by focus effect on the target segments.

Table 2.5 The results of post hoc test for fricative duration

	Real word		Nonsense word	
/s/	IP, PP IP, Wd PP, Wd	p=.162 p=.007 p=.003	IP, PP IP, Wd PP, Wd	p=.105 p=.015 p=.027
/s'/	IP, PP IP, Wd PP, Wd	p=.032 p=.060 p=.736	IP, PP IP, Wd PP, Wd	p=.026 p=.030 p=.115

2.3.2.1.2 CV Duration

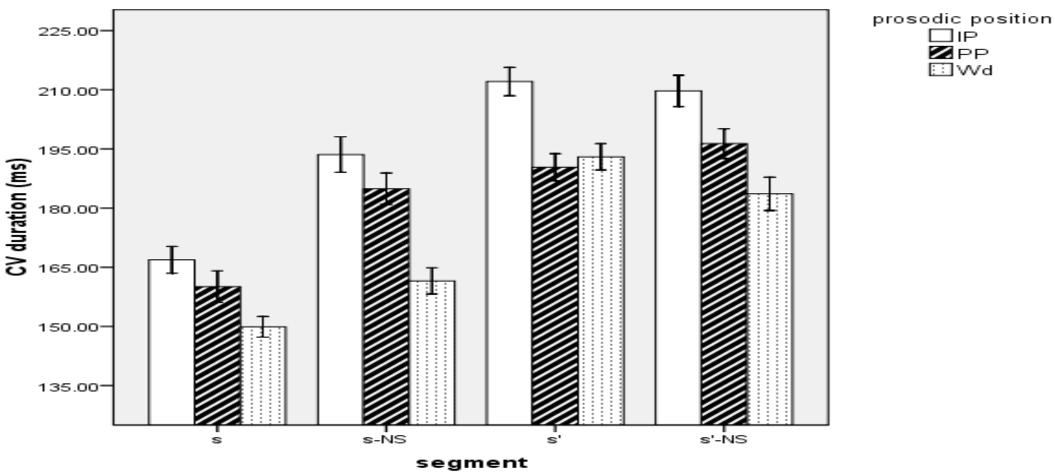


Figure 2.8 Pooled graph for CV duration by segment type (combined with word type) * prosodic position

In figure 2.8, CV durations in real and nonsense words were compared at three different prosodic positions. The CV durations of /s/ and /s’/ are longest in IP initial position, intermediate in PP-initial position, and shortest in Wd-initial position, except for /s’/ in real words. For /s’/, there is no difference between PP and Wd-initial positions in real words. The results of CV durations show the same pattern with the results of fricative duration.

The speakers M1, M2 and F1 in Appendix A.2.(2) displayed the same pattern with the pooled graph. However, the other speakers did not follow this pattern and the

CV duration in IP-initial position was not longer than in the other prosodic positions. Compared to the results of fricative duration, there was some inconsistency across speakers in the results of CV durations.

The main effect for fricative category was significant but the other main effects were not significantly different ($F(1, 5) = 37.874$, $p = .002$ for fricative category; $F(2, 10) = 1.779$, $p = .218$ for prosodic position; $F(1, 5) = .487$, $p = .516$ for word type). The interaction of fricative category, prosodic position and word type was significant ($F(2, 10) = 6.953$, $p = .013$**).**

The results of Post-hoc comparisons are presented in Table 2.6. For /s/ in the real words, the CV duration did not significantly differ across each prosodic domain-initial position. For /s/ in nonsense words, the CV durations were significantly different between IP and Wd, and between PP and Wd-initial positions but there was no significant difference between IP and PP-initial positions. For /s'/, there was significant difference between IP and PP in real and nonsense words, and between IP and Wd in nonsense words. In the comparison between real and nonsense words, only lenis fricatives showed significant difference between the two word types in IP ($p=.015$), PP ($=.018$), and Wd-initial positions ($p < .000$). However, fortis fricatives did not show a significant difference between the two word types. The two fricatives show different prosodic position effect on CV duration and the effect of word type depends on the effect of fricative category.

Table 2.6 The results of post hoc test for CV duration

	Real word		Nonsense word	
/s/	IP, PP IP, Wd PP, Wd	p=.516 p=.138 p=.080	IP, PP IP, Wd PP, Wd	p=.078 p=.023 p=.019
/s'/	IP, PP IP, Wd PP, Wd	p=.012 p=.067 p=.192	IP, PP IP, Wd PP, Wd	p=.007 p=.011 p=.069

2.3.2.1.3 Final Vowel Duration

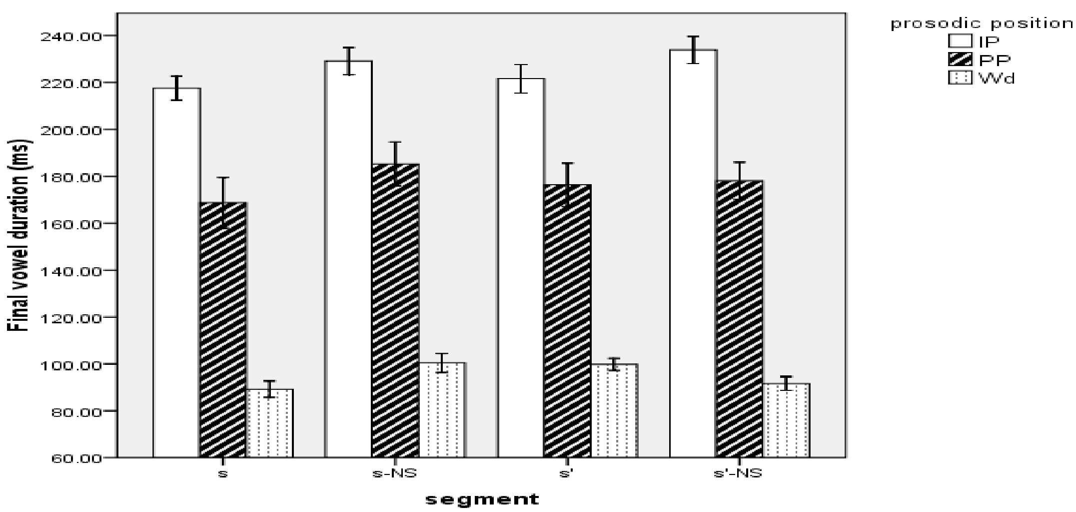


Figure 2.9 Pooled graph for final vowel duration by segment type (combined with word type) * prosodic position

Figure 2.9 displays vowel durations before /s/ and /s'/ in real and nonsense words across IP, PP and Wd-final positions. The pooled graph illustrates that the final vowel is longest in IP, intermediate in PP and shortest in Wd. However, the graphs of individual speakers in Appendix A.2.(3) show that the speakers M2 and M3 have the longest vowel duration in PP-final position.

The main effect for prosodic position was significant but the effects of fricative category and word type were not significant. ($F(1, 5) = .902$, $p = .386$ for fricative category; $F(2, 10) = 22.994$, $p < .000$ for prosodic position; $F(1, 5) = .001$, $p = .981$ for word type). The interaction of three factors was not significant ($p = .627$). But, the interaction of fricative category and prosodic position was significant ($p = .029$) and the interaction of prosodic position and word type was highly significant ($p < .000$). The significant interaction seems to be due to the highly significant effect of prosodic position in final vowel duration.

Table 2.7 illustrates the results from Pairwise post hoc tests for final vowel duration. There were significant differences of vowel duration between IP and Wd, and between PP and Wd-final positions both in real and nonsense words, except for /s'/ in nonsense words. Before /s'/ in nonsense words, the vowel duration was significantly distinctive across the three prosodic domains. But in general, the final vowel duration showed the pattern IP, PP >>Wd. There was less consistent distinction between IP and PP due to longer final vowel duration in PP produced from speaker M2 and M3.

Table 2.7 The results of post hoc tests for final vowel duration

	Real word	Nonsense word
/s/	IP, PP: p=.161 IP,Wd: p<.000 PP,Wd: p=.022	IP, PP: p=.083 IP,Wd: p<.000 PP,Wd: p=.012
/s'/	IP, PP: p=.119 IP,Wd: p<.000 PP,Wd: p=.017	IP, PP: p=.036 IP,Wd: p <.000 PP,Wd: p=.006

2.3.2.2 Spectral Parameters

2.3.2.2.1 Centroid Frequency

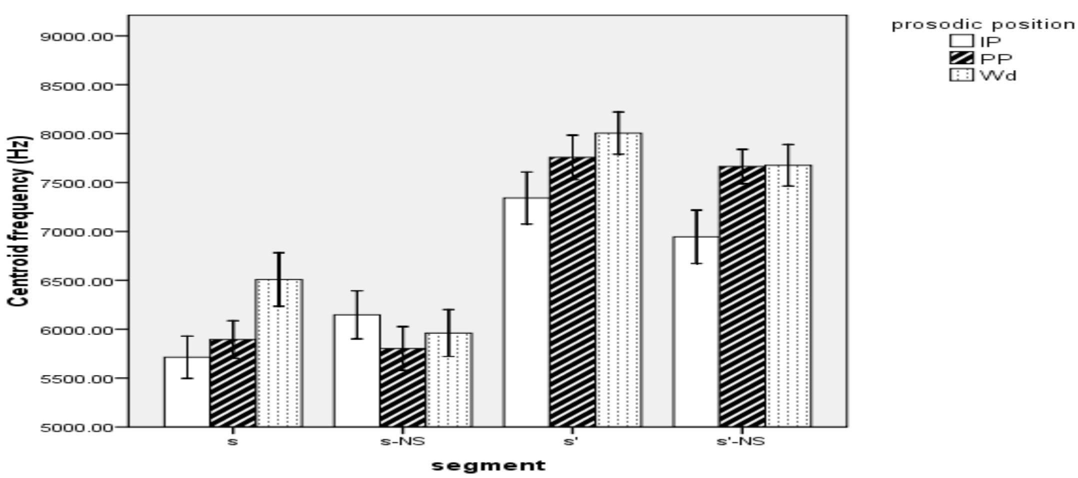


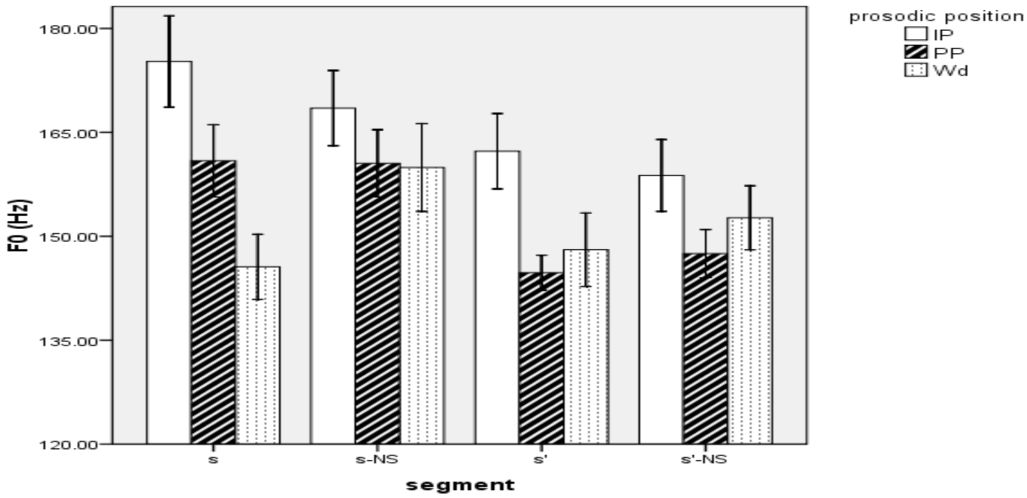
Figure 2.10 Pooled graph for centroid frequency by segment type (combined with word type) * prosodic position

In Figure 2.10, centroid frequency of each fricative is compared at IP, PP and Wd-initial positions in real and nonsense words. The pooled graph shows that the centroid frequencies of /s/ and /s'/ in real words are progressively higher when the prosodic domain gets lower. But, in nonsense words, /s/ does not show a systematic pattern across prosodic domains, and /s'/ has lower centroid frequency in IP than in the other prosodic domains while there is no difference between PP and Wd-initial positions.

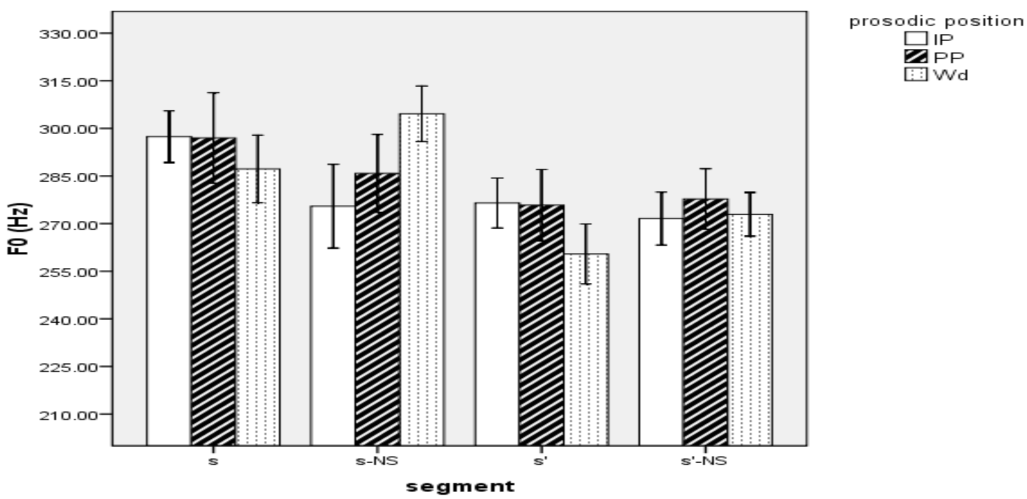
The main effect for fricative category was significant but the other main effects were not significant ($F(1, 5) = 11.663$, $p = .019$ for fricative category; $F(2, 10) = 1.572$, $p = .255$ for prosodic position; $F(1, 5) = .965$, $p = .371$ for word type). None of the interactions were significant. Since the graphs from individual speakers in the Appendix A.2.(4) did not show the same pattern as in the pooled graph and there was no statistically significant difference among the three prosodic positions, the result did not support the hypothesis in that centroid frequency is higher in higher prosodic positions than in lower prosodic positions.

Since the centroid frequency of /s'/ was shown to be greater than that of /s/ in all prosodic positions, this results support Cho et al. (2002)'s claim that, in Korean, fortis fricatives are produced with a relatively smaller front cavity.

2.3.2.2.2 Fundamental Frequency



a. Pooled graph from male speakers



b. Pooled graph from female speakers

Figure 2.11 Pooled graph for F0 by segment type (combined with word type) * prosodic position

Figure 2.11 presents F0 after each fricative across prosodic domain-initial positions both in real and nonsense words. The pooled graph from male speakers in Figure 2.11a shows that F0 is highest in IP-initial position among three different prosodic positions but there is no systematic difference between PP and Wd-initial positions. The pooled graph from female speakers in Figure 2.11b does not show any consistent difference across three prosodic positions.

The main effect for fricative category was significant but the other main effects were not significant ($F(1, 5) = 22.141$, $p = .005$ for fricative category; $F(2, 10) = 2.178$, $p = .164$ for prosodic position; $F(1, 5) = 1.436$, $p = .284$ for word type). None of the interactions were significant. In contrast to the results of F0 after stop categories, there is no word type effect in the results of F0 after fricative categories. The significant effect of fricative category was due to the higher F0 after lenis fricatives than after fortis fricatives. In Cho et al. (2002), no significant difference of F0 was found between the two fricative categories in the Seoul speakers. But since the target fricatives were produced in isolation, it is hard to compare the results of F0 to the findings in this study. In addition, subjects tend to put focus on the target words although the focus positions of test sentences were controlled. Because of random focus, the results of F0 do not seem to show systematic pattern in this study.

2.3.2.2.3. [H1-H2]

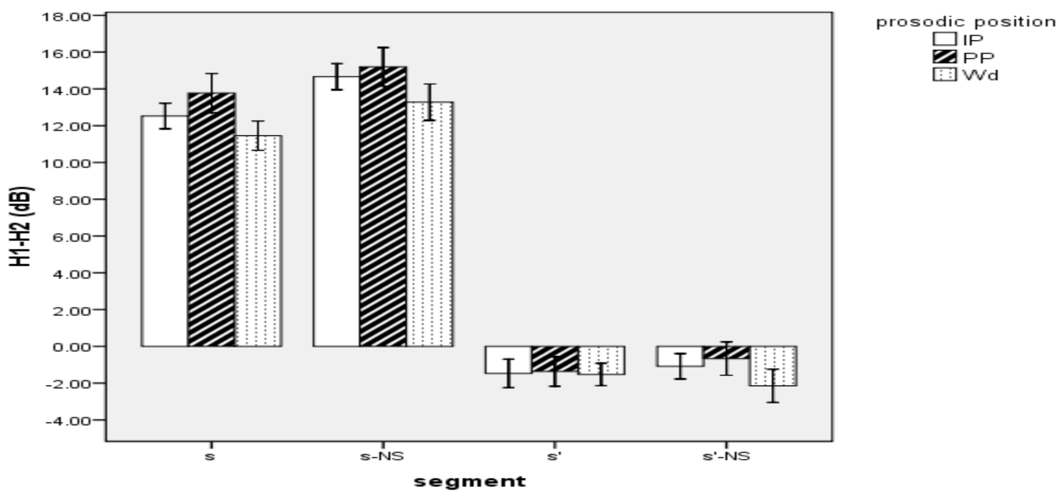


Figure 2.12 Pooled graph for [H1-H2] by segment type (combined with word type) * prosodic position

In Figure 2.12, [H1-H2] difference after fricatives was compared at IP, PP and Wd initial positions in real and nonsense words. It is seen that there is no systematic distinction among the three prosodic domain-initial positions, but that [H1-H2] values are different between the two fricatives. The graphs of the individual speakers in Appendix A. 2.(6) show that all speakers produced greater [H1-H2] after lenis fricatives than after fortis fricatives. For fortis fricatives, one male and all female speakers have negative values in [H1-H2]. They produced more creaky vowels after fortis fricatives than the other two male speakers.

The main effect for fricative category was significant but the other main effects were not significant ($F (1, 5) = 94.667$, $p < .000$ for fricative category; $F (2, 10) = .726$, $p = .508$ for prosodic position; $F (1, 5) = 2.595$, $p = .168$ for word type). None of the interactions was significant**.** The significant effect of fricative category was caused by the different [H1-H2] between /s/ and /s'/. This result is compatible with the results of Cho et al. (2002).

S. Kim (2001) reported that one of the two speakers showed a significant difference in [H1-H2] between IPi and APm (AP medial position), and between APi and APm. Since all six speakers in this study showed different patterns across prosodic domains in the results of [H1-H2], it is hard to conclude that there is significant prosodic effect on [H1-H2].

2.4 Discussion

The results of various analyses of prosodic domain-initial effects confirmed the fact that Korean stops and fricatives have a limited set of acoustic parameters that show more enhanced acoustic properties in higher prosodic domains relative to lower prosodic domains. The significant prosodic domain-initial effects were found in the results of duration intervals. The results of this study support the hypotheses that VOT for stops, fricative duration for fricatives, and CV duration for both stops and fricatives are greater in higher prosodic domain-initial positions than in lower prosodic domain-initial positions.

With regard to the variability depending on phonation and articulation types, as noted in Cho & Keating (2001), Korean fortis stops did not show variation in VOT across prosodic domain-initial positions. However, lenis and aspirated stops showed enhanced durations as a function of prosodic domains. Due to a greater variation of lenis stops, VOT values for lenis stops in IP and PP were greater than those for aspirated stops in Wd domains. So, in the higher prosodic domain-initial positions, the contrast between fortis stop and the other stops is enhanced but the contrast between lenis and aspirated stops decreases. For fricatives, both fricatives categories showed enhanced fricative duration and CV duration in higher prosodic domain-initial positions but there was less variation across prosodic domain-initial positions for fortis fricatives relative to lenis fricatives.

The prosodically-driven properties in domain-initial positions were also affected by word types (real vs. nonsense words). The results of VOT, fricative duration and CV duration support the hypothesis that those durations are longer in nonsense words than in real words. That is, the duration parameters in prosodic domain-initial positions were more enhanced in nonsense words than in real words. However, a significant durational difference between the two word types was only found in lenis stops and fricatives, as aspirated stops, and fortis stops and fricatives did not show enhanced duration in nonsense words. In Kang & Guion (2008), they found that the VOT difference between lenis and aspirated stops was enhanced in clear speech and that the VOTs for aspirated stops were relatively stable across speaking style. They explained that since stimuli were elicited in the utterance-initial position, the high VOT values for aspirated stops were already enhanced by domain-initial strengthening and may be close to the phonetic target. Likewise the aspirated stops in the current study were enhanced in acoustic duration at the initial position of higher prosodic domains and might have overshot the phonetic target. As a result, the increased duration was not shown in nonsense words.

Another reason for less variability in those phonation types is related to their articulatory properties. Cho & Keating (2001) claimed that fortis and aspirated stops can be considered to be a “strong” consonant type because they have greater linguopalatal contact compared to nasal and plain stops. Aspirated and fortis stops were known to have longer acoustic closure duration, more tongue blade contact, higher tongue movements and higher glottal raising than lenis coronal stops. In MRI, acoustic and aerodynamic study of fricatives (Kim et al. (2005)), it was found that contrary to /s/, /s’/ has longer and narrower oral constriction, greater pharyngeal width and the highest tongue blade and glottal height was sustained longer. Less variability in articulation could be a reason for less variability of duration in prosodic position and word type effects for fortis and aspirated stops, and fortis fricatives. On the other hand, lenis categories seem to have more than enough room for variation in linguopalatal contact and acoustic duration as a function of prosodic position and word type.

In the results of RMS burst energy for stops, centroid frequency for fricatives, and F0 and [H1-H2] of the following vowels for stops and fricatives, there was no significant prosodic domain-initial effect. Since segment effects were found for these spectral properties, these results reflect the distinctions among phonation types for stop and fricative categories. The results of F0 showed significant word type effect for stops but there was no significant word type effect for fricatives. In short, since the prosodic domain-initial effect was not found in the properties of the vowel following stops and fricatives, the phonetic properties correlated with prosodic structure are limited to the durational parameters of initial consonant in the prosodic domains in Korean.

When it comes to the results of F0, male speakers produced higher F0 after stops and fricatives in IP than in the other prosodic domains. But female speakers did not show this pattern for stops and fricatives. However, it does not seem to be caused by gender difference. Rather, it was due to random focus on target segments since each speaker showed different pattern in the results of F0. Although focus was controlled with wh-questions and focus-cueing sentences, speakers put focus on target segments and this might cause inconsistent results in F0.

With regard to the distinction across different prosodic domains, not all prosodic domains were distinguished by the acoustic properties analyzed in the current study. For example, VOT and CV durations for lenis and aspirated stops were not significantly different between IP and PP-initial positions, but showed a significant difference between IP and Wd, and PP and Wd. So, the results of VOT and CV durations showed the pattern, IP, PP >> Wd. In the results of fricative durations, the lenis fricatives revealed the pattern IP, PP >> Wd but fortis fricatives did not show the pattern. The results of CV durations for fricatives showed the pattern, IP, PP >> Wd for lenis fricatives and IP >> PP, Wd for fortis fricatives in nonsense words. Furthermore, the distinction of prosodic domains was not consistent between the two word types. In general, there was a robust distinction between IP and Wd-initial positions but the distinction between these and with the intermediate phrase was relatively variable. A summary of the distinctions as a function of prosodic levels is shown in Table 2.8.

The distinction of prosodic domains was also variable in the results of final vowel durations as shown in Table 2.9. It has already been noted that final vowel lengthening is strong evidence for a prosodic domain boundary. As in the results of prosodic domain-initial properties, final vowel durations showed distinction between IP and Wd levels but the inconsistent distinction of the two higher prosodic domains (IP and PP) was also found in the results of final vowel durations.

Table 2.8 Summary of distinction in prosodic domain-initial positions

	Real word	Nonsense word
VOT for /t, t^h/	IP, PP >> Wd	IP >> PP >> Wd for /t/ IP, PP >> Wd for /t^h/
CV duration for /t, t^h/	IP, PP >> Wd	
Fricative duration for /s, s'/	IP, PP >> Wd for /s/ IP >>PP for /s'/	IP, PP >> Wd for /s/ IP >>PP, Wd for /s'/
CV duration for /s/	IP >> PP for /s'/	IP, PP >> Wd for /s/ IP >> PP, Wd for /s'/

Table 2.9 Summary of distinction in prosodic domain-final positions

	Real word	Nonsense word
Lenis stops	IP >> PP >> Wd	IP >> PP >> Wd
Aspirated stops	IP, PP >> Wd	IP, PP >> Wd
Fortis stops	IP >> PP >> Wd	IP >> PP >> Wd
Lenis fricatives	IP, PP >> Wd	IP, PP >> Wd
Fortis fricatives	IP, PP >> Wd	IP >> PP >> Wd

Cho & Keating (2001) found a four-way distinction among prosodic positions (Ui > IPi > APi > Wi). However, they did not control focus in their study and the test sentences in their stimuli were relatively simple and had many repetitions (20 repetitions) for articulatory analysis. It caused speakers to put focus on the target segments and to have more distinctions across different prosodic positions. These factors might explain the difference between their results and the results of the current study.

One of the other reasons for different distinctions across prosodic domains is speaker variation. For example, speaker F3 had longer VOT in PP than in IP for /t/ in real words and no difference of VOT between IP and PP for /t^h/ in real and nonsense words as shown in Appendix A.1.(1) This speaker also showed a similar pattern for fricative durations. For /s/ in real and nonsense words, fricative duration was longer in PP-initial position than in IP-initial position. This speaker variation might wash out the distinction between IP and PP in the pooled results.

In addition to the prosodically driven properties in stops and fricatives, this study gives us several previously unknown properties. Jun (1993) and Kang (1992) suggested that intervocalic voicing applies to lenis stops at the medial position of AP or PP but not at the initial position of the intermediate phrase in Korean. The percentage of voicing of the entire stop closure duration in Cho & Keating (2001) showed the pattern, IPi < APi < Wi. Lenis stops had 100 % closure voicing at AP-medial (Wd-initial) but the voicing varied for AP-initial lenis stops. However, in the current study, lenis stops between vowels did not show voicing even in Wd-initial position (PP-medial position) except for one speaker. Jun (1993) has already demonstrated that Lenis stop voicing is a gradient phonetic rule since the rule is sensitive to speech rate and phrasing, segmental and prosodic contexts. The findings in the current study might be reflections of hypercorrection in the linguistically artificial environment of the phonetics laboratory. However, since the voicing rule showed variation across different studies, the grammatical status of Lenis stop voicing rule needs to be revised.

When it comes to the results in [H1-H2] after fortis category, it has been found that fortis stops and fricatives showed very small or negative [H1-H2] values (Abberton (1972), Ahn (1999), and Cho et al. (2002)). But the results of the current study revealed that not all speakers showed negative [H1-H2] values as in the pooled graph in Appendix A.1.(6) and Appendix A.2.(6). All female speakers and only one male speaker showed negative [H1-H2] values after fortis category. The other two male speakers showed very small positive [H1-H2] values for the fortis category compared to the lenis and aspirated categories. Female speakers showed higher F0 than male speakers, and the male speaker M2 had a relatively high f0 after target segments than the other male speakers as shown in Appendix A.1.(5) and Appendix A.2.(5). Since the higher F0 is, the farther apart the harmonics are, there might be more difference of energy between H1 and H2. So, it seems that the subjects who have relatively higher f0 tend to produce more distinction in [H1-H2] values across different phonation types. However, in order to support this idea, a more quantitative experiment needs to follow.

In conclusion, as predicted, the left edge of prosodic domain is marked by enhanced durational properties for Korean stops and fricatives. As shown in the previous studies, the prosodic domain-initial properties are cumulative when the prosodic domain gets higher. The prosodically-driven durational properties are also affected by different word type (real vs. nonsense words). The initial properties across different prosodic domains showed more enhanced properties in nonsense words than in real words. In other words, the segmental properties are affected by information content, in terms of prosodic position and word type. When there is less contextual information as in prosodic domain-initial position and in nonsense words, the enhanced properties of initial segments may provide perceptual cues for the higher information content of the segment. But it was also found that these effects vary depending on phonation types in Korean. Not all segments showed enhanced acoustic properties in the initial position of prosodic domains. Different enhancing strategies among phonation types would maximize phonological contrast and the different variation across prosodic positions could be perceptually distinguished.

Chapter 3: The Perception of Korean Stops and Fricatives in IP, PP and Wd-initial Positions

The results of the production study in the previous chapter demonstrated that Korean stops and fricatives have longer durations (VOT, fricative duration, CV duration) at higher prosodic domain-initial positions than at lower prosodic domain-initial positions. When comparing two word types (real and nonsense words), the durational properties in prosodic domain-initial positions were more enhanced in nonsense words than in real words. The prosodically-driven phonetic properties were found to be distinct across phonation types in stop and fricative categories. However, it is not known whether listeners are sensitive to these enhanced properties driven by prosodic position and whether the enhanced properties in nonsense words were perceptually distinguished. The second experiment is a perception experiment to test the perceptual distinctiveness of mismatched CVs from different prosodic domains.

3.1 Objectives

The objective of the perception study is to test my main hypothesis that the prosodically conditioned realizations of segments will be perceptually distinguished. I examined whether cross-spliced CVs from different prosodic domains (i.e., when a target

CV is placed in IP-initial position, the CV is taken from a segmentally matched token in PP and Wd-initial positions) affect the perception of target segments in continuous speech. I compared the perceptual difference between the cross-spliced CVs and same-spliced CVs (i.e., when the target CV is placed in IP-initial position, the CV is extracted from a segmentally matched token in another IP-initial position). In addition, I examined whether the specific acoustic properties of the different articulations and phonation types uncovered in the production study affect the perception of these target segments.

In the perception study, I expect that error rates in identifying the cross-spliced CVs will be significantly greater than the same-spliced CVs. Because of the different prosodically conditioned properties, subjects are expected to have more difficulty in identifying cross-spliced CVs in continuous speech. This difficulty in identification could also be reflected in reaction time. So, I expect that listeners will identify the same-spliced CVs more accurately and faster than cross-spliced CVs.

In the perception of cross-spliced CVs, I expect that subjects will make significantly more errors identifying CVs extracted from the lower prosodic domains than from the higher prosodic domains. Because of the relatively shorter durations and less distinctive phonological contrast, listeners are expected to have more difficulty in identifying the target segments from lower prosodic domain-initial positions.

Due to the great durational variation of lenis stops in higher prosodic domain-initial positions, the phonological contrast between lenis and aspirated stops decreased and perceptual confusions between the two phonation types can occur in the conditions.

So, in the perception of lenis stops, I expect that subjects will make more errors in identifying a CV from the higher prosodic domains when the CV is placed in lower prosodic domains. Only for lenis stops, the target CVs from both higher and lower prosodic domains are expected to affect the identification of lenis stops. On the contrary, fortis stops did not show variation in durational parameters across prosodic domain-initial positions. So, the identification of fortis stops is expected to be constant when the underlying prosodic position of the stop is altered.

Lastly, when there is less top-down processing effect as with nonsense words, it is expected that subjects will make significantly more errors identifying the CV in nonsense words than in real words. Ganong (1980) reported that in identifying /t/ or /d/ on the basis of VOT, subjects tend to identify the segment as one that forms a real word, where auditory information is ambiguous (i.e., VOT is less informative). So, I expect that auditory information will be relied on more heavily when there is no lexical effect. In addition to the effect of top-down processing, the enhanced properties in nonsense words are expected to affect the perception of the target segments. When the acoustic properties of segments in real words were compared to those in nonsense words, lenis stops and fricatives showed significant statistical differences between the two word types in durational parameters. The increased duration for lenis stops caused less distinctive contrast between lenis and aspirated stops in nonsense words. In addition, the enhanced distinction of phonetic properties across prosodic domain-initial positions might give rise to more confusion in the identification of target segments. So, I expect that prosodic domain-initial segments in nonsense words will be misidentified more often than those in real words.

The proposed hypotheses are summarized as in (5).

(5) Research Hypotheses

a. In the perception of cross-spliced CVs, subjects will make more errors in identifying CVs from the lower prosodic domain-initial positions than from the higher prosodic domain-initial positions.

b. In the identification of lenis stops, subjects will make more errors in identifying CVs from the higher prosodic domain-initial positions than from the lower prosodic domain-initial positions.

c. Error rates in identifying cross-spliced CVs will be greater than same spliced CVs.

d. The reaction time in identifying same-spliced CVs will be faster than cross-spliced CVs.

e. Subjects will make significantly greater errors identifying CVs from nonsense words than from real words.

3.2 Methods

3.2.1 Original Stimuli

Recordings from male speaker (M1) in the production experiment were used in this study. The male speaker is a graduate student at the University of Texas at Austin who had lived in Seoul until age 30.

The phonation type of the initial segment of a CV was varied across aspirated, lenis and fortis stops (e.g., /t^h/, /t/, /t'/), and lenis and fortis fricatives (e.g., /s/, /s'/) in Korean. The target CV was placed at IP, PP and Wd-initial positions both in real and nonsense words.

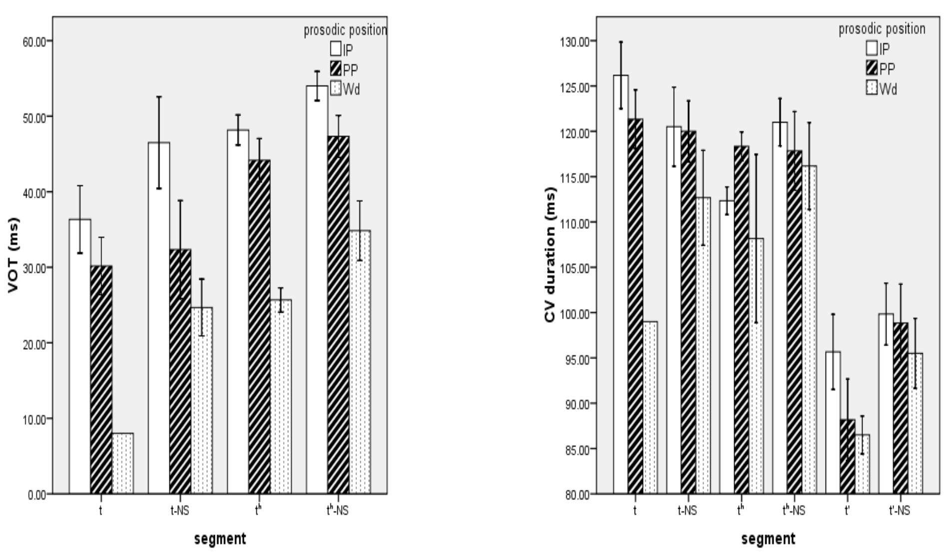
I briefly summarize the results of production experiment from M1. Figure 3.1 displays the results of VOT, CV duration, relative RMS, F0, [H1-H2] and final vowel duration for stops. Because /t'/ did not show difference as a function of prosodic position, it was not included in the figure of VOT as in Figure 3.1. The results of VOT for /t/ and /t^h/ showed a progressively increasing duration in the higher prosodic domain-initial positions. Except for /t^h/ in real words, the increasing trend was also found in the results of CV duration. Final vowels before stops showed longer duration in higher prosodic domains than in lower prosodic domains. In the spectral parameters, the relative RMS burst energy was greater in lower prosodic domain-initial positions than in higher prosodic domain-initial positions. F0 was higher in the higher prosodic domain-initial positions than in lower prosodic domain-initial positions. But, the results of [H1-H2] did not show consistent pattern across different prosodic domains.

The table 3.1 summarizes the results from M1 with a three way ANOVA. Relative to the results in the pooled data, the speaker M1 shows significant effects of segment type and prosodic position in all measurements except for the results of relative RMS burst energy. The effect of word type is significant in VOT, CV duration and [H1-H2].

Figure 3.1 The results of production study for stops from speaker M1[18]

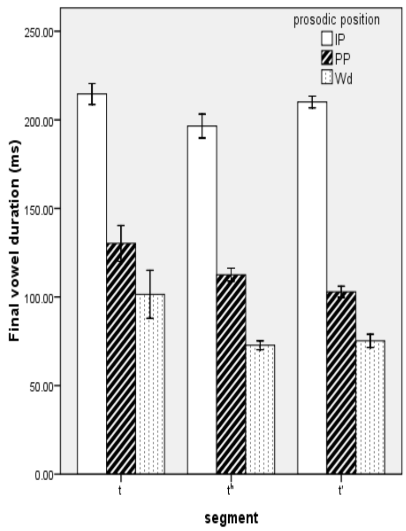
(i) Durational Parameters

a. VOT b. CV duration



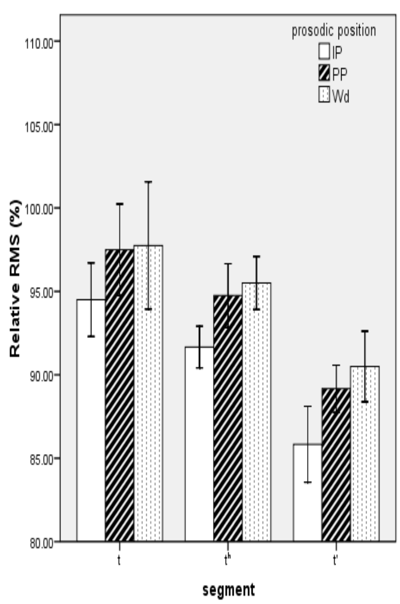
[18] Since significant word type effect was only found in the results of VOT, CV duration and [H1-H2], the results from nonsense words were not separately presented in the graphs of the other results.

c. Final vowel duration

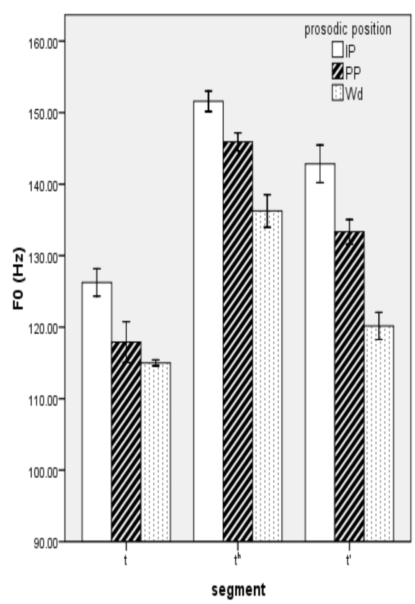


(ii) Spectral Parameters

d. Relative RMS burst energy



e. F0



f. H1-H2

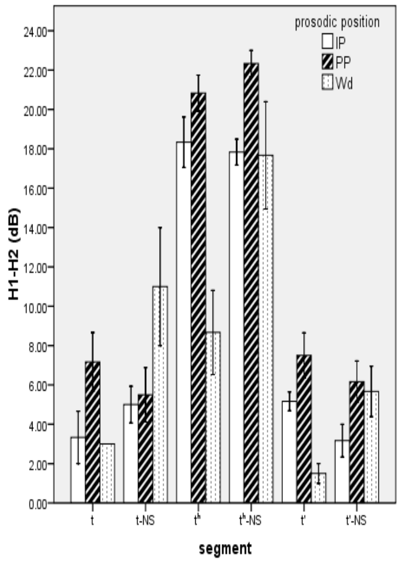


Table 3.1 The results of three-way ANOVA for stops from speaker M1

	segment type	prosodic position	word type	Interaction
VOT	p < .000	p <.000	p=.002	Segment * prosodic position (p <.000)
CV duration	p < .000	p =.006	p=.025	
Relative RMS	p < .000	p =.090	p=.851	
F0	p < .000	p < .000	p=.547	Prosodic position* word type (p=.003)
H1-H2	p <.000	p<.000	p=.005	Segment*prosodic position (p=.004) Prosodic position *word type (p<.000)
Final vowel duration	p = .002	p < .000	p=.785	

Fig. 3.2 displays the results of the production experiment for fricatives from M1. The fricative and CV duration show an upward trend as a position moves up in the prosodic hierarchy. The final vowel was also lengthened in the higher prosodic domains. In spectral parameters, the results of F0 also showed an increasing trend but the results of [H1-H2] and centroid frequency did not show consistent variation across prosodic positions.

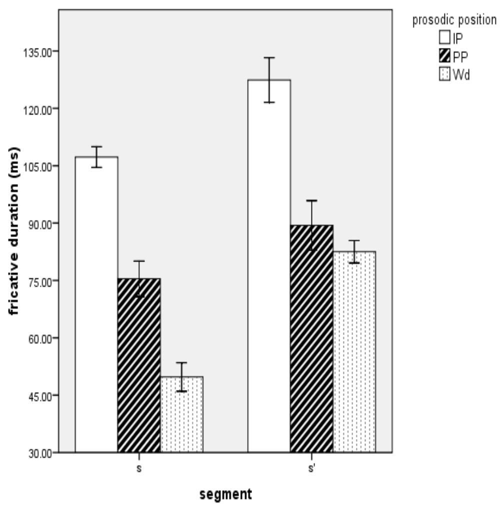
Table 3.2 summarizes the results of the three-way ANOVA for fricatives from speaker M1. There were significant effects for segment type and prosodic position for all properties but the word type effect was significant in the results of centroid frequency and [H1-H2].

To summarize, the results of M1 exhibited a significant prosodic effect in both durational and spectral parameters for stops and fricatives, except for the results of relative RMS burst energy. These results differ from the statistical results of the pooled data. With regard to word type effect, the results from M1 showed significant word type effect in VOT and CV duration for stops, as seen in the pooled data, but did not show significant word type effect in fricative duration and CV duration for fricatives, thus contrasting with the results of the pooled data.

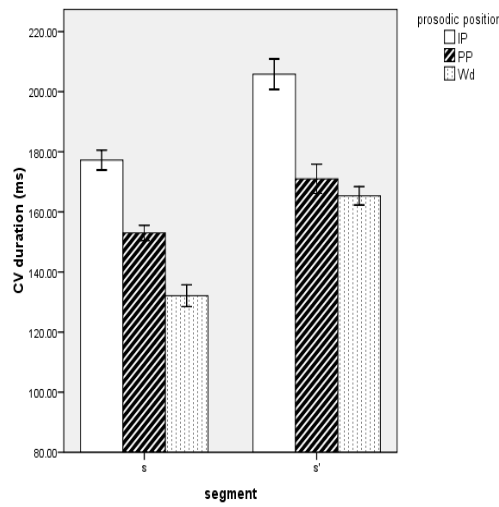
Figure 3.2 The results of production study for fricatives from speaker M1[19]

(i) Durational Parameters

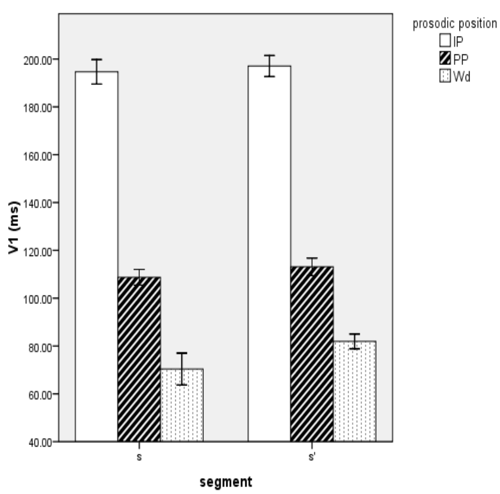
a. Fricative duration



b. CV duration



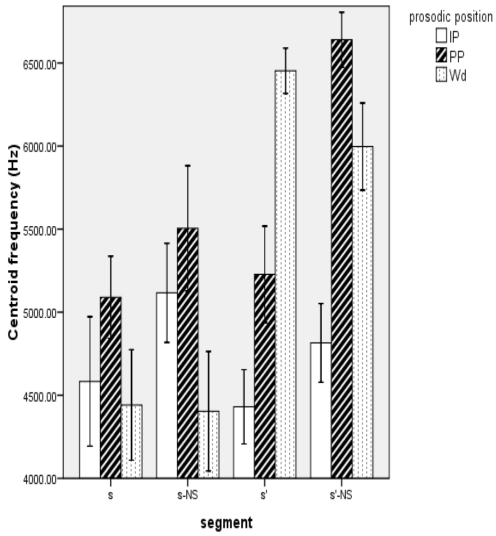
c. Final vowel duration



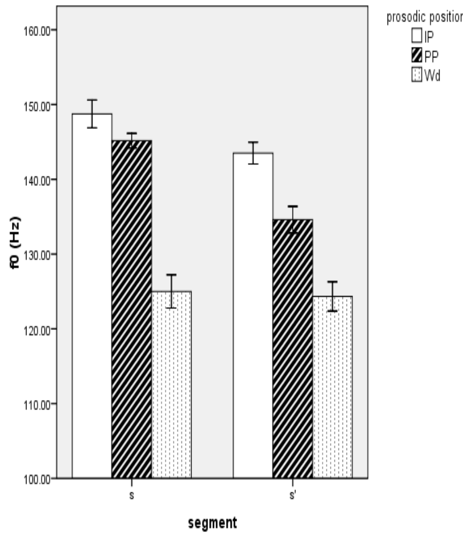
[19] The results from nonsense words were separately presented in the figure of centroid frequency and [H1-H2] since they showed significant word type effect.

(ii) Spectral Parameters

d. Centroid frequency



e. F0



f. H1-H2

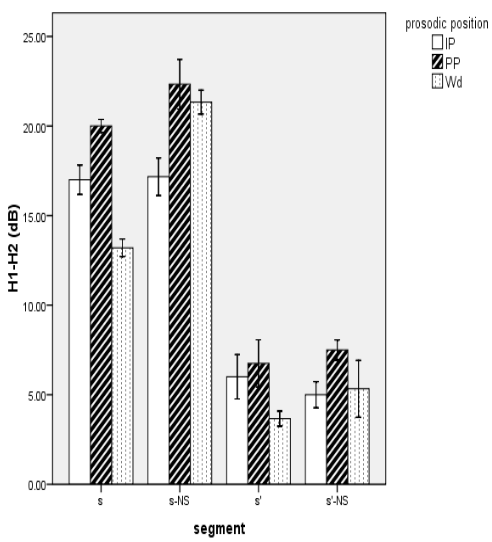


Table 3.2 The results of three-way ANOVA for fricatives from speaker M1

	Effect of segment type	Effect of prosodic position	Effect of word type	Interaction
Fricative duration	$p < .000$	$p < .000$	$p = .275$	
CV duration	$p < .000$	$p < .000$	$p = .870$	
Centroid frequency	$p < .000$	$p < .000$	$p = .029$	Segment * prosodic position ($p < .000$)
F0	$p < .000$	$p < .000$	$p = .130$	Prosodic position* word type ($p < .000$)
H1-H2	$p < .000$	$p < .000$	$p = .001$	Segment*prosodic position ($p = .012$) Prosodic position *word type ($p = .002$)
V1	$p = .105$	$p < .000$	$p = .485$	

3.2.2 Manipulated Stimuli

The initial CV was chosen for splicing, rather than just the onset consonant because the effect of strengthening was not limited to the initial segment, but was found to spread into the following segment (Cho (2004), Cho et al. (2007)). In addition, since the stop and fricative categories in Korean are in part differentiated by the properties of the following vowel (Kim et al. (2002)), the initial CV makes the target word more natural in continuous speech. For the stops, to avoid the pause included in the IP boundary, the duration of the silent stop closure was not included in the splicing. So, the onset of each CV was defined as the release of the closure for stops and the beginning of the high frequency fricative noise for fricatives. The offset of each CV was defined as the offset of the vocalic energy associated with the following vowel /a/ of the target segment. All splices were made at zero-crossings in the speech waveform by using Praat. The sample display with splicing point is shown in Fig. 3.3.

Each stop (or fricative)-plus-vowel sequence was cross-spliced from the portion of another waveform containing the identical segmental context produced at a different prosodic level. For example, in IP-initial position, /ta/ was spliced from the portions of the two waveforms containing /ta/ in PP-initial and Wd-initial positions, respectively. Next, in PP-initial position, /ta/ was extracted from the portions of the two waveforms including /ta/ in IP and Wd-initial positions. Finally, in Wd-initial position, /ta/ was taken from /ta/ in IP and PP-initial positions, respectively.

To make CV sequences that remain in their original prosodic position, a CV sequence was spliced from the identical segmental token at the same prosodic position. This procedure was meant to ensure that the splicing technique is the same for all of the stimuli. These procedures produce 9 different prosodic conditions as shown in Table 3.3 The shaded cells indicate same-spliced prosodic conditions.

This identical procedure was repeated for real words and nonsense words. The same splicing procedure was applied to the filler sentences including affricates, /ʧa/,

/ʧʰa/ and /ʧ'a/ as well. I used Praat in splicing the target CV in the different prosodic domains.

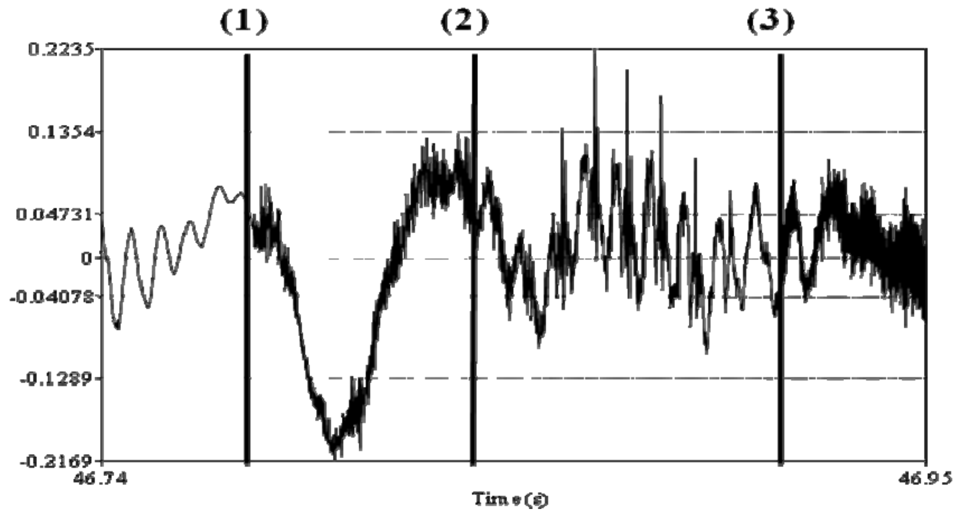
Table 3.3 9 prosodic conditions from same and cross-splicing

The prosodic position in the utterance	The prosodic position where CV is spliced from		
/ta/$_{IP}$	/ta/$_{IP\text{-}ip}$	/ta/$_{IP\text{-}pp}$	/ta/$_{IP\text{-}wd}$[20]
/ta/$_{PP}$	/ta/$_{PP\text{-}ip}$	/ta/$_{PP\text{-}pp}$	/ta/$_{PP\text{-}wd}$
/ta/$_{Wd}$	/ta/$_{Wd\text{-}ip}$	/ta/$_{Wd\text{-}pp}$	/ta/$_{Wd\text{-}wd}$

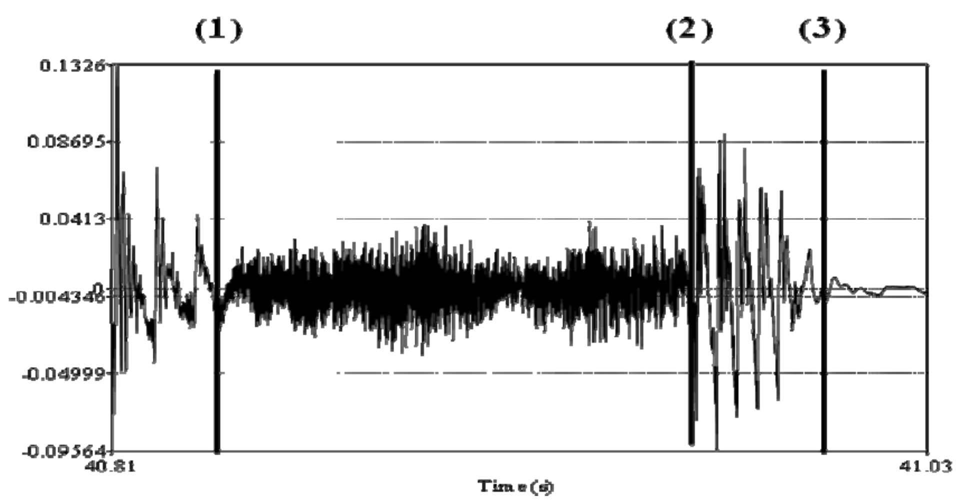
[20] The capital letters represent the prosodic position of the CV in the carrier utterance and the following subscript letters refer to the position where the CV is spliced from. For example, in IP-wd, the target CV is placed in Intonational phrase-initial position of utterance but it is spliced from Word-initial position.

Figure 3.3 Waveforms for one token of the original lenis stop and fricative stimuli. The three vertical lines indicate, from left to right, consonant onset (1), consonant offset/vowel onset (2), and vowel offset (3).

a. Stops



b. fricatives



3.2.3 Listeners

10 female and 10 male listeners participated in the perception study. They were phonetically untrained native speakers of the Seoul dialect of Korean. Their ages ranged from 25 to 40 years. Most were graduate or undergraduate students of the University of Texas at Austin or were otherwise affiliated with the university. Their length of residence in the U.S. ranged from 5 months to 8 years. None of the speakers have any known speech or hearing disorders.

3.2.4 Procedures

The randomized test stimuli were presented using the Alvin program on an IBM laptop computer. Participants were tested individually in a quiet room. It took approximately 35 minutes per subject. They were told that they will hear different sentences containing the initial syllable such as /ta/, /tʰa/, /t’a/, /sa/, /s’a/, /ʧa/, /ʧʰa/ and /ʧ’a/ in the middle of test sentences. For each one they were asked to select the closest sound from predesigned answers on the screen written in Korean orthography. Response errors and reaction time were collected. All participants received practice trials before the experiment.

3.2.5 Statistical Design

There were 5 target segments (/t/, /t^ha/, /t'a/, /sa/ and /s'a/), 9 prosodic conditions (IP-ip, IP-pp, IP-wd, PP-ip, PP-pp, PP-wd, Wd-ip, Wd-pp and Wd-wd) and 2 word types (real and nonsense words). Four sentences were tested for each condition and 54 filler sentences were included. Overall, a total of 8280 identifications were collected ((5 × 9×2×4 +54) × 20).

Mean error rates and reaction time for same and cross-spliced CVs were measured in the identification test. A repeated measures analyses of variance (RM ANOVAs) was performed with three within-subjects factors (segment type, prosodic position, and word type) using SPSS 1.6.

3.3 Results

3.3.1 Identification Error Rates for Korean Stops

In the results using error rates as the dependent variable, there were highly significant main effects for segment type ($F(4, 76) = 8.580$, $p < .000$), prosodic condition ($F(8, 152) = 19.053$, $p < .000$) and word type ($F(1, 19) = 100.366$, $p < .000$). The interaction of segment type, prosodic condition and word type was also highly significant (p<.000).

Figure 3.4a, b, and c display error rates for the 9 different prosodic conditions in real and nonsense words for lenis, aspirated and fortis stops. Overall, subjects made more errors in identifying the target CVs in nonsense words than in real words. The error rates for lenis stops were greater than the other stop categories in all prosodic conditions.

For the lenis stop /t/, when the target CV was cross-spliced to IP and PP-initial positions from Wd-initial position (IP-wd, PP-wd), listeners made greater errors in identifying the CV in real and nonsense words. The maximal distinction of phonetic properties between IP and Wd caused the highest error rates on IP-wd condition among all prosodic conditions.

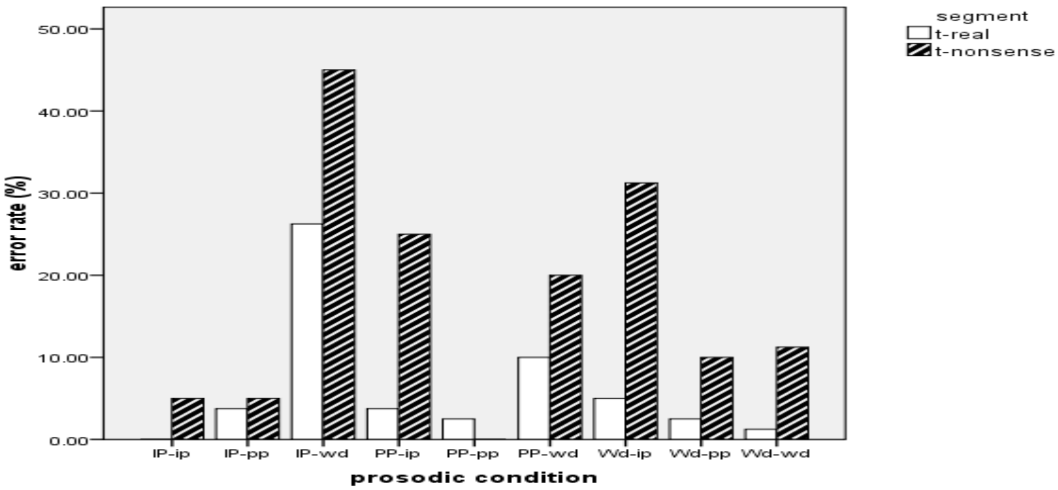
When the CV was extracted from IP-initial position, subjects made more errors in identifying the CVs in PP and Wd-initial positions (PP-ip, Wd-ip). The enhanced VOT and CV duration in the higher prosodic domains caused less distinctive contrast between lenis and aspirated stops, and it gave rise to confusion in the identification of lenis stops in lower prosodic domain-initial positions. Although the significant difference between IP and PP-initial positions was not found in the results of VOT and CV duration, the relatively longer durations from IP-initial position affected the identification of target CVs in PP.

For the aspirated stop /t^h/, subjects showed higher error rates in nonsense words relative to very small error rates in real words. When the target CV was placed in IP and

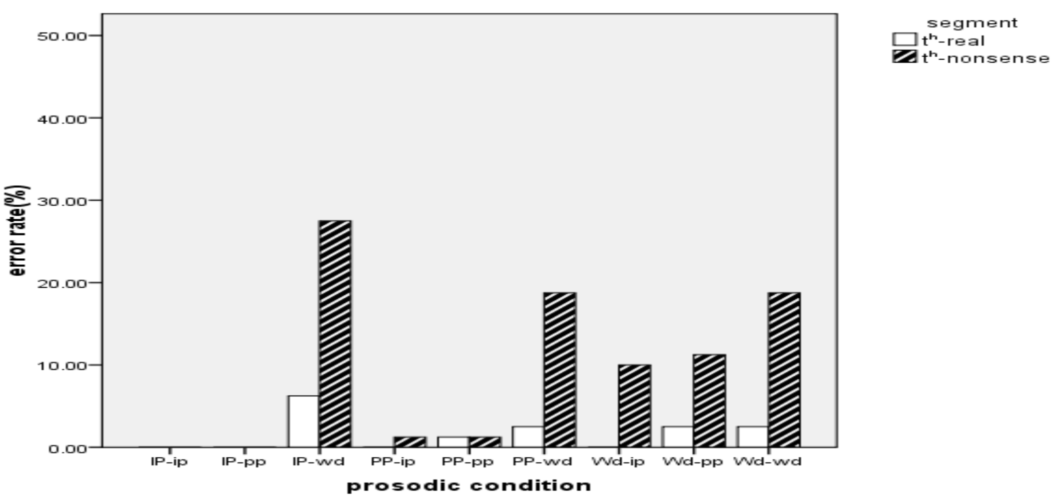
PP-initial positions, listeners made more errors in identifying the CV extracted from Wd-initial position. The CVs extracted from Wd-initial position also caused more errors in the identification of aspirated stop CVs when they are placed in Wd-initial position. The results of VOT revealed that VOT values of lenis stops in IP-initial position were overlapped with those of aspirated stops in Wd-initial position. The relatively short duration and less distinctive phonological contrast in Wd level affected the perception of both same and cross-spliced CVs taken from Wd domain. However, subjects did not make more errors in identifying CVs for aspirated stops from higher prosodic domain-initial positions.

For fortis stops, listeners did not make many errors in identifying the target CVs from different prosodic positions and they showed less than 10 % error rates for 9 prosodic conditions. Contrary to the error rates for lenis and aspirated stops, there were no significant differences between real and nonsense words and between same-spliced CVs and cross-spliced CVs.

a. Error rates for /t/



b. Error rates for /t^h/



c. Error rates for /t'/

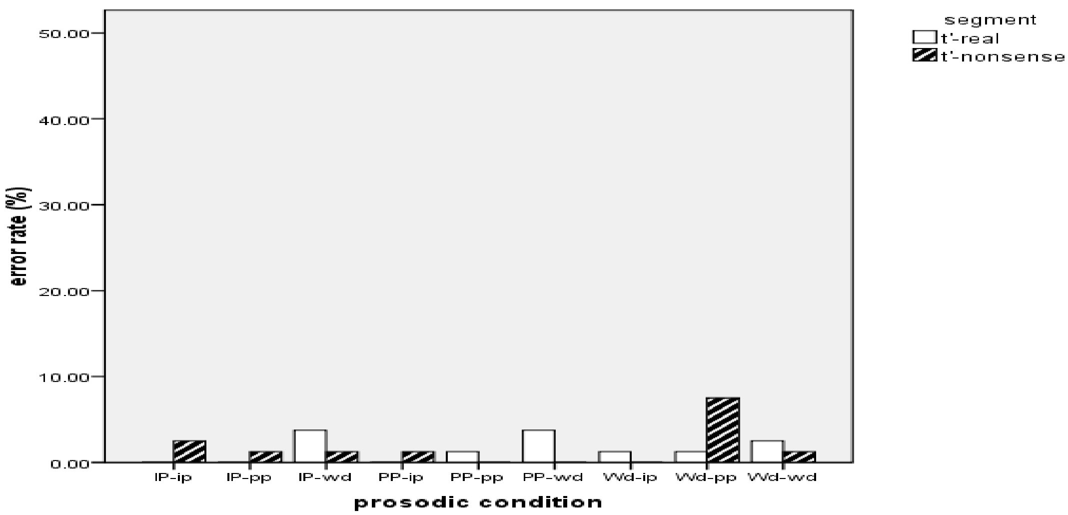


Figure 3.4 Identification error rates for stops

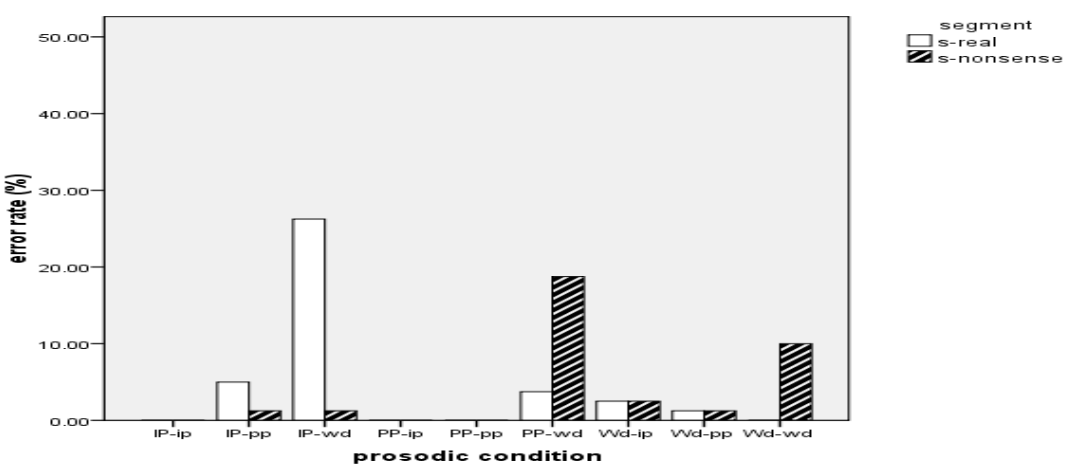
3.3.2 Identification Error Rates for Korean Fricatives

The identification error rates for lenis and fortis fricatives are presented in Figure 3.5a, b. Contrary to the results for stops, the identification error rates for fricatives were relatively small. Overall, listeners showed more error rates in lenis fricatives than in fortis fricatives.

For lenis fricatives /s/, when the target CV was placed in IP-initial position, subjects made more errors in identifying the CVs extracted from Wd-initial position in real words. When the CV was placed in PP and Wd-initial positions, listeners made more errors in identifying the CV from Wd-initial position in nonsense words. That is, across three different prosodic domains, subjects made more errors in identifying the target CVs when the CV was spliced from Wd-initial position. Similar to the results of aspirated stops, there was no increase in error rates in identifying the CV from higher prosodic domains. Contrary to the lenis and aspirated stops, there is no consistent word type effect in the results of error rates for lenis fricatives.

For fortis fricatives, subjects did not make many errors in identifying the target CV from all 9 prosodic conditions. All error rates were less than 5 %. The significantly different properties found in fricative duration and CV duration across three prosodic domains did not influence the perception of same and cross-spliced CV of fortis fricatives. The results of identification error rates for fortis fricative were similar to those for fortis stops. There were no consistent patterns across prosodic positions and between word types.

a. Error rates for lenis fricative



b. Error rates for fortis fricative

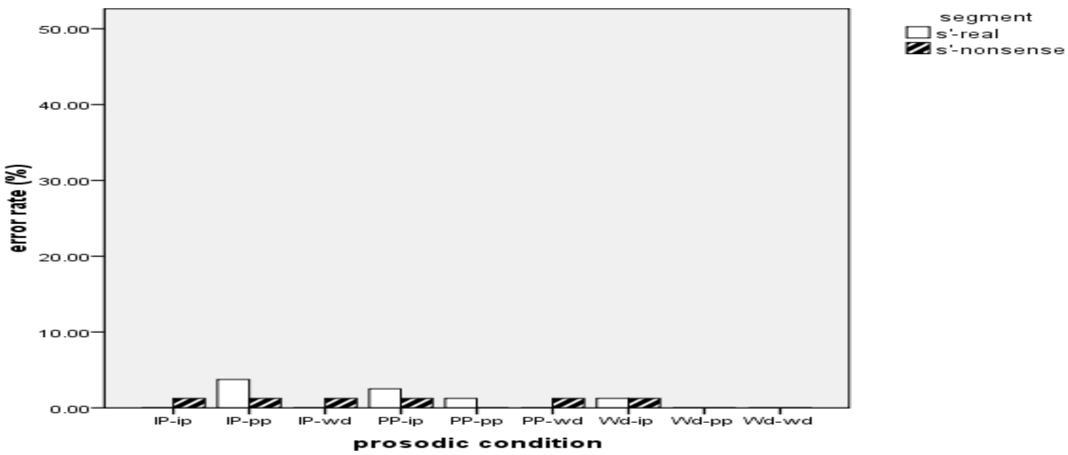
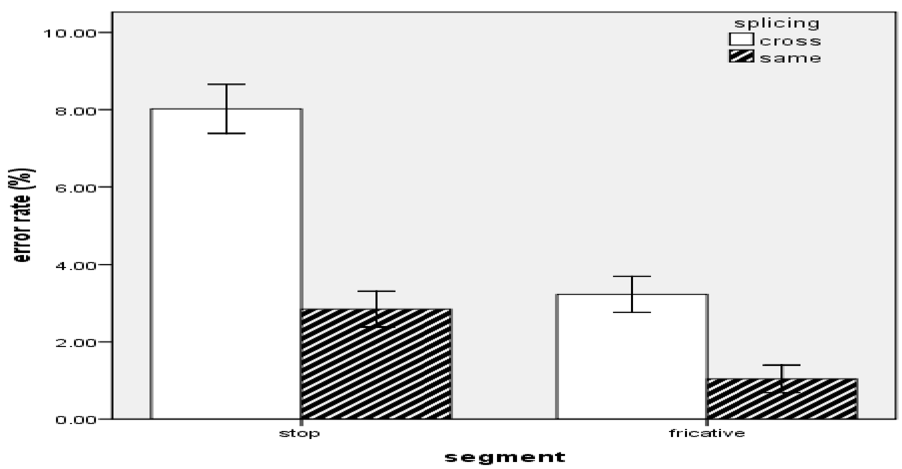


Figure 3.5 Identification error rates for fricatives

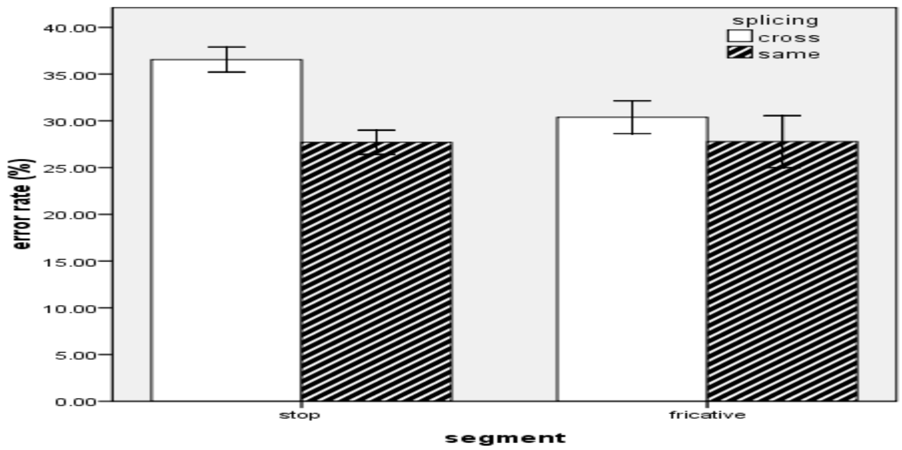
3.3.3 Error Rates for Same-spliced CVs vs. Cross-spliced CVs

Figure 3.6a displays identification error rates between same and cross-spliced CVs for stops and fricatives. Subjects made more errors in identifying the cross-spliced CV than the same-spliced CV for both stops and fricatives. In addition, the error rates in identifying same and cross-spliced CVs were greater for stops than for fricatives. However, because of greater percentage of correct answers (which was measured as zero), the mean error rates for all conditions in Figure 3.6a were very small. In order to compare only misidentified tokens between same and cross-spliced CVs, the data were filtered to include only error rates over zero and the results are shown in Figure 3.6b. The comparison of misidentified tokens between the two conditions also revealed that the error rates in cross-spliced CVs were greater than those in same-spliced CVs.

In order to find out the perceptual difference between same and cross-spliced CVs, planned comparisons were performed. The results showed highly significant differences between same-spliced CVs and cross-spliced CVs ($p < .000$) and between the two segments ($p < .000$). As predicted, subjects made significantly more errors identifying cross-spliced CVs than same-spliced CVs. In addition, the results displayed that listeners made smaller error rates between same and cross-spliced CVs for fricatives than for stops.



a. Identification error rate on same and cross-spliced CVs



b. Identification error rate on same and cross-spliced CVs with filtered tokens (error rates over zero)

Figure 3.6 Identification error rates on same and cross-spliced CVs

3.3.4 Reaction Time Results for Same-spliced CVs vs. Cross-spliced CVs

To compare the reaction time between same and cross-spliced CVs, the reaction times from all subjects were collected except for one subject. This subject took 20 minutes more than the rest of the subjects in the identification test, distorting the overall pattern in the results. So, I excluded the reaction time from this subject. The results of reaction time in Fig. 3.7 show that listeners were faster to respond to same-spliced CVs than to cross-spliced CVs: 381 vs. 397 ms on average for stops, 393 vs. 407 ms on average for fricatives (F (1, 18) = 13.582, p =.002 for segment; F (1, 18) =

32.309, p <.000 for splicing). So, the results in the reaction time provide evidence supporting the hypothesis that subjects have more difficulty in identifying cross-spliced CVs due to the different prosodically conditioned properties.

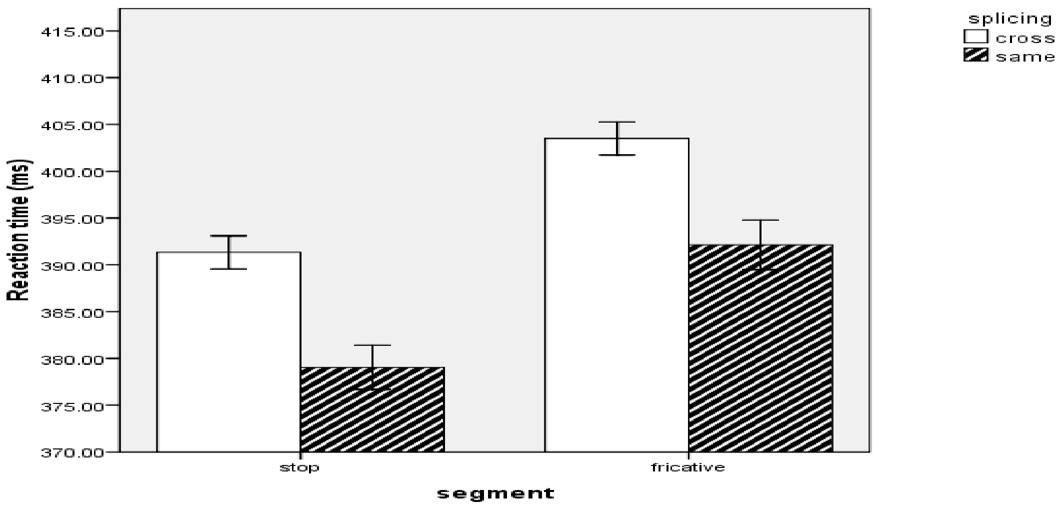


Figure 3.7 Reaction time for same and cross-spliced CVs

3.4 Discussion

The results of the perception experiment reveal that Korean listeners are sensitive to the prosodically driven properties of target CV segments in the perception of continuous speech. The results from identification error rates and reaction times for same and cross-spliced CVs support hypotheses (5a) and (5b) in that listeners perceived the same-spliced CVs more accurately and faster than cross-spliced CVs.

In addition, the hypothesis in (5c) was supported in that subjects made more errors in identifying the CVs from lower prosodic domain-initial position than from higher prosodic domain-initial position. For example, the error rates for IP-wd were greatest in all prosodic conditions for lenis stops (approximately 50%) and for aspirated stops (approximately 40%) and lenis fricatives (approximately 40%). When the target CVs were placed in PP-initial position, subjects also made greater errors in identifying the CV extracted from Wd-initial position. The acoustic results of VOT, fricative duration and CV durations were significantly shorter in Wd-initial position than in the other higher prosodic domain-initial positions. It might cause less distinctive phonological contrast in lower prosodic domain-initial positions. As a result, subjects had difficulty in identifying the target segments from lower prosodic domain-initial position. The distinctive acoustic properties of each prosodic domain-initial position were perceptually distinguished by Korean listeners.

The greater variability across prosodic domains for lenis stops also affected the perception of the segments in different prosodic domains. Lenis stops had significantly longer VOT intervals in higher prosodic domain-initial positions and the values were close to those of aspirated stops, resulting in reduced phonological contrast between lenis and aspirated stops. As a result, in addition to the misidentification of lenis stops from lower prosodic domain-initial positions, subjects made more errors in identifying the CVs from higher prosodic domains. This pattern only occurred in the perception of lenis stops.

The most interesting finding of this experiment was seen in the comparison of error rates between real and nonsense words. The results from the production study showed significant word type effects in the acoustic properties for VOT, fricative duration and CV duration in lenis stops and fricatives, and for F0 in stops. But the enhanced acoustic properties did not cause enhanced intelligibility and contrast in nonsense word condition. There were higher error rates for lenis and aspirated stops and lenis fricatives in nonsense words than in real words.

One of possible reasons for higher error rates in nonsense words is that subjects seem to rely on the auditory information more than the lexical information in identifying nonsense words. It has been known that context can control the perception of individual segments (Warren (1970) and the identification of sequences of words was aided when the words form sequences in continuous speech. However, when there is no lexical effect, it can be expected that listeners are more sensitive to the distinctive acoustic properties of the segments. Perhaps due to less contextual information, the acoustic differences across prosodic domains were more easily detected in the identification of target segments in nonsense words. Another possible reason for the result is that there is less phonological contrast between categories in nonsense words. For example, the contrast between lenis and aspirated stops was found to be less distinctive in nonsense words than in real words due to the greater durational variation of lenis stops relative to that of aspirated stops. Less phonological contrast might have caused more confusion between the two segments in the identification tests.

However, for fortis fricatives, the acoustic variation found in the production study was not reflected in the perception study. Contrary to fortis stops, fortis fricatives showed longer fricative duration and CV duration in IP-initial position than in PP and Wd-initial positions although the durational variation as a function of prosodic position was less consistent for fortis fricatives compared to lenis fricatives. However, like the results for fortis stops, the identification error rates for fortis fricatives were extremely small for all prosodic conditions. In other words, variation of durational parameters for fortis fricatives did not cause confusion with the other category. Since both fortis stops and fricatives showed relatively very small error rates compared to the other phonation types, further research is required to determine what the strong perceptual cues are independent of prosodic conditions.

The conclusions drawn from the results of the perception study are that the prosodically driven phonetic properties across prosodic levels are perceptually distinctive. The enhanced acoustic properties of higher prosodic domain-initial segments were shown to be less confusable compared to those of lower prosodic domain-initial segments, except for lenis stops. It provides evidence that the enhanced acoustic properties give perceptual cues for initial segments with higher information content.

However, since this study was performed with natural continuous speech, it cannot be determined yet what specific acoustic cues played a role in triggering the listener's identification response. The test stimuli contain prosodic domain-initial properties as well as phrase final properties such as final vowel lengthening and boundary tone in each prosodic domain. The suprasegmental properties of target CVs might also affect the identification of altered CVs. So, to discover the salient acoustic cues that potentially influence the perception of the prosodic domain-initial segments, further speech editing studies are needed.

Chapter 4: Conclusions

4.1 Summary of the Study

This study investigated the production and perception of Korean stops and fricatives as a function of their position in prosodic hierarchical domains.

The results of the production study confirmed the fact that both Korean stops and fricatives exhibited enhanced phonetic properties in the initial position of higher prosodic domains. The duration intervals of initial segments were cumulatively lengthened for lenis and aspirated stops and lenis fricatives as a position moves up in the prosodic hierarchy. These results are in accord with findings in previous studies (Cho & Keating (2001), S. Kim (2001)).

The prosodic domain-initial effects were also affected by the information content of words. The durational intervals for lenis stops and fricatives in prosodic domain-initial positions were longer in nonsense words than in real words. On the other hand, fortis stops did not show any variation as a function of prosodic position and word type and fortis fricatives showed less variation compared to lenis fricatives. The significant findings are summarized in Table 4.1. However, the results from spectral parameters did not support the hypotheses in that relative RMS burst energy for stops and centroid frequency for fricatives were not consistently influenced by prosodic positions and word types. Furthermore, the prosodic domain-initial effect did not extend to vowel quality (F0 and [H1-H2]). Only for stops, F0 was shown to be significantly higher in nonsense words than in real words.

Table 4.1 Summary of ANOVA results for durational variation

	VOT			CV duration			Fricative duration		CV duration	
	/t/	/t^h/	/t'/	/t/	/t^h/	/t'/	/s/	/s'/	/s/	/s'/
Prosodic position effect	*	*	ns.	*	*	ns.	*	*	*	*
Word type effect	*	ns.	ns.	*	ns.	ns.	*	ns.	*	ns.

* indicates significant results at 0.05, and ns. indicates non-significant results.

The perception experiment sought to determine if there was a functional role for prosodically driven phonetic properties on the perceptual identification of Korean stops and fricatives. The results of the current study revealed that Korean listeners are sensitive to the prosodically driven phonetic properties from different prosodic domain-initial positions in the perception of continuous speech.

As expected, listeners identified the same-spliced CVs more accurately and faster than cross-spliced CVs. Since the pre-boundary properties (lengthening or boundary tone) were not manipulated in this study, the results confirmed the fact that the prosodic domain-initial properties from different prosodic domains influence the perception of the target CVs.

For the identification of cross-spliced CVs, the CVs taken from lower prosodic domains resulted in greater error rates in the identification of the target segments when placed in higher prosodic domains. For lenis stops, listeners also had higher error rates in identifying CVs extracted from IP-initial position when the target CVs were placed in PP and Wd-initial positions. Overall, due to the enhanced distinction across prosodic domains and decreased top-down processing effect in nonsense words, the identification error rates were greater in nonsense words than in real words.

When target segments did not show greater durational enhancement in prosodic domain-initial positions due to less variability in their articulation, the identification error rates for the segments were negligible. For example, less variability of fortis stops and fricatives across prosodic domains was also reflected in less misidentification of those segments. The lack of articulatory adjustments due to prosodic domain position was correlated with highly consistent perceptual performance.

4.2 Enhanced Phonological Contrast as a Function of Prosodic Position and Word Type

Due to different prosodic domain-initial effect on durational parameters, the phonological contrast for phonation types was more enhanced in higher prosodic positions and in nonsense words than in lower prosodic positions and in real words. For example, the results from VOT showed that the contrast between lenis and aspirated stops, and fortis stops was more enhanced in higher prosodic domains and in nonsense words. However, the contrast between lenis and aspirated stops becomes less distinctive since the duration parameters for lenis stops increased to a great degree in the higher prosodic positions and in nonsense words but aspirated stops showed slightly greater VOT values in those conditions as shown in Figure 4.1. The mean VOT for lenis stops in nonsense words was very close to the mean VOT for aspirated stops, while VOT for lenis stops in real words was comparatively lower than that for aspirated stops.

a. VOT contrast for real words

b. VOT contrast for nonsense words

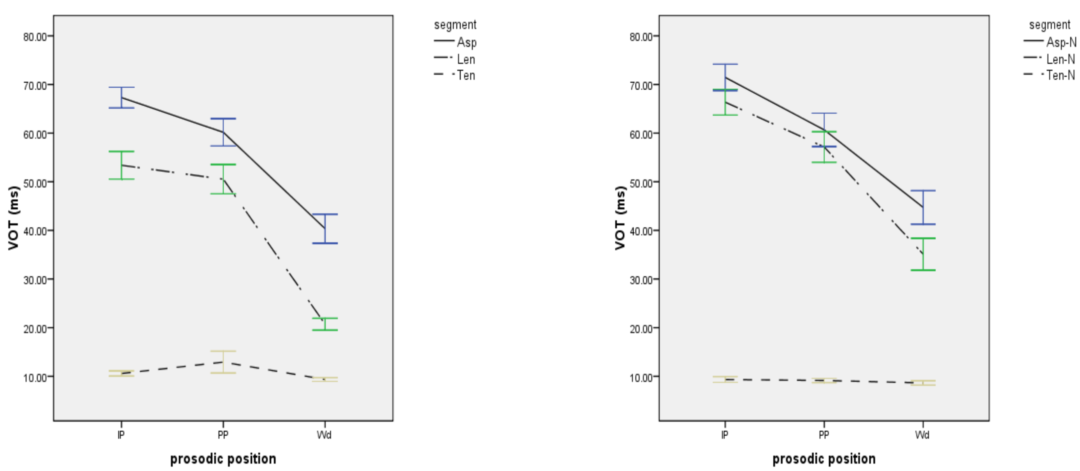


Figure 4.1 Variation of VOT as a function of prosodic position and word type

Kang & Guion (2008) demonstrated a similar pattern of enhanced phonological contrast in Korean stops through comparisons of different speech styles. The aspirated and lenis stops contrast was found to be expanded in clear speech by enhancing VOT for older speakers and F0 for younger speakers. Their results showed that VOT for lenis stops decreases in clear speech, resulting in enhanced contrast between lenis and aspirated stops.

In the study of domain-initial effects on bilabial stops /p, p^h, p'/, Cho & Jun (2000) found that the three stop categories were maximally dispersed for VOT and integrated airflow in IPi position. They claimed that laryngeal features are enhanced domain initially. Following Lombardi's privative (i.e. unary) feature system, the increased VOT and airflow for the prosodic domain-initial /p^h/ were explained as enhancement of the feature [spread glottis], and the invariant VOT of /p'/ was described as enhancement of the feature [constricted glottis]. Since lenis stops /p/ were unspecified for both features, they explained that lenis stops augmented syntagmatic CV contrast[21] (contrast between the initial segment and its neighbors).

However, in comparison of VOT between lenis and aspirated stops, the results of the current study showed different findings. As shown in Figure 4.1, the VOT contrast between lenis and aspirated stops becomes less distinctive in higher prosodic domain and nonsense words. Cho & Jun (2000) demonstrated that the variation for lenis stops rarely overlapped with that of aspirated stops, and the enhanced properties did not blur the

[21] Cho & Jun (2000:2) noted that the most commonly agreed characteristic of the domain-initial effect is to enhance 'consonantality' of the segment, resulting in the syntagmatic contrast with the following vowel.

contrast between aspirated and lenis stops. But in the current study, VOT values for lenis stops in IP and PP-initial positions did overlap with those for aspirated stops in Wd-initial positions. In comparison of glottal opening among stop categories in Kim et al. (2009), word-initial /p/ and /k/ were also found to have greater glottal opening than word-medial /p^h/ and /k^h/.

Due to the different place of articulation of target stops and relatively short test stimuli with many repetitions (9 times), the results in Cho & Jun (2000) seemed to show more distinctive articulations among three different stops. Furthermore, since lenis stops showed much variation in VOT across positions and word types and even different studies, the classification with unspecified features for lenis stops seems problematic as an explanation for the variation.

In a study of acoustic and auditory correlates of vowels in focused context, Hay et al. (2006) suggested that the amount of contrast varies greatly between non-focused condition and focused condition. For example, duration ratios between German long/short vowel pairs were found to be larger than those between English fortis/lax vowel pairs in non-focused condition, and only German talkers increased vowel duration ratio in focused condition. These results are in part analogous to the findings of this study. The amount of distinction between aspirated/lenis stops and fortis stops in Wd domain and real words was much enhanced in higher prosodic domains and in nonsense words. It is interesting that the increased VOT enhanced the distinction between the stops with aspiration and stops without aspiration in lower prosodic domains and real words. These results also imply that the greater the phonological contrast between the categories, the more they are perceived as distinct, whereas the less contrast between the categories, the more confusion between them. In other words, listeners are expected to perceive the distinctive phoneme better in the position with higher information content as in higher prosodic domain-initial positions.

For example, Figure 4.2 displays error rates for cross-spliced CVs of Korean stops. Since the phonological contrast was relatively less distinctive in lower prosodic positions than in higher prosodic positions, listeners had difficulty in identifying the CVs extracted from lower prosodic domain-initial positions. In the perception study, it was found that the robust distinctions between the highest constituent IP and the lowest constituents Wd were consistently reflected in the identification accuracy of the target segments. The error rates for IP-wd were greatest for all segments except fortis stops. The less distinctive contrast between lenis and aspirated stops in higher prosodic domain was also reflected in the error rates for PP-ip and Wd-ip conditions.

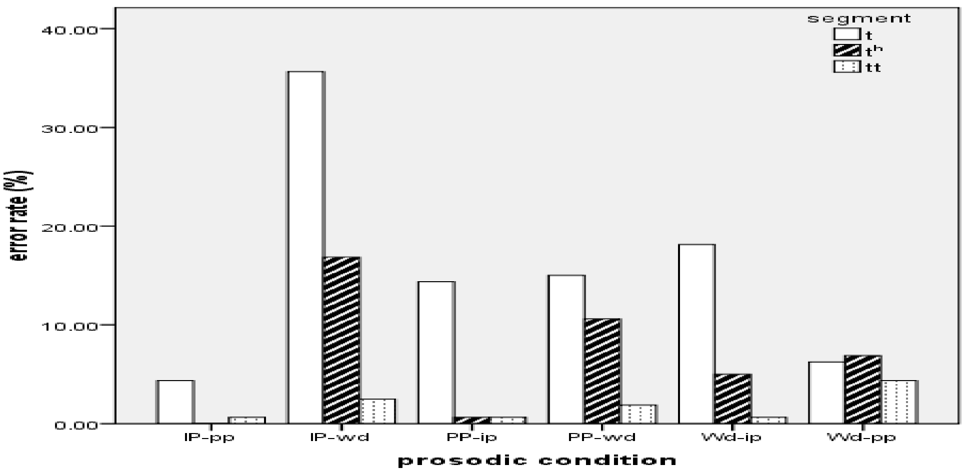


Figure 4.2 Identification error rates for cross-spliced CVs of stops

When it comes to the contrast between fricative categories, it is plausible to assume that the amount of distinction of fricative categories in Wd level was greatly enhanced in higher prosodic domains. However, the results of durational parameters for fricatives did not show the contrastive distinction for fricative categories as stops did. There was even no significant difference in the fricative duration between /s/ and /s'/.

Fricatives are known to be less subject to articulatory variation in general because they tend to be more constrained in their articulatory and acoustic properties (Fougeron (2001)). Acoustically, their inherent longer duration with intense noise might cause less distinctive durational enhancement across prosodic domains compared to stops. This may cause less prosodically dependent variation in the production and perceptual confusion of fricatives in identification tests.

In addition, as noted in Chapter 1, it was found that contrary to /s'/, /s/ is produced with greater glottal opening and has aspiration interval in word-initial position (Kagaya (1974), Jun et al. (1998), Cho et al. (2002)). Yoon (1999) found that before mid and low vowels, the duration of the aspiration interval was the only consistent difference between the two Korean fricatives. On the other hand, Chang (2007) reported that segmental duration was not a significant perceptual cue but rather that the combination of F1 onset, intensity buildup and voice quality were the most important perceptual cues for Korean fricative categories. Aspiration duration was found to have a strong effect on perception of /s/ in addition to vocalic cues (a high F1 onset, gradual intensity buildup and positive [H1-H2]). However, in the current study, the vocalic information and the duration of the aspiration interval has not been manipulated in the spliced CVs. Thus, it may cause relatively small identification error rates for fricatives.

To be brief, the results of fricatives do not directly support the fact that the contrast in Wd level of real speech was enhanced in higher positions. But since the listeners had difficulty in identifying /s/ when they were taken from Wd-initial position, it implies that there was more confusion in the identification of /s/ in the lower prosodic domain-initial positions than in the higher prosodic domain-initial positions.

In order to find out whether the less distinctive acoustic properties and contrast between categories caused confusion in the identification response, confusion matrices are presented as in Table 4.2. In this table, subject's response types are listed horizontally at the top of each matrix and target segments are shown vertically. Since Korean affricates (/ʧ/, /ʧ^{h}/, /ʧ’/) were included as filler utterances, they were also listed in the identification test. The “none” columns refer to the answers showing none of the listed segment types.

Table 4.2 Confusion matrices for initial segments spliced from IP, PP and Wd domains.

a. CVs spliced from IP-initial position

	t	t^{h}	t’	s	s’	ʧ	ʧ^{h}	ʧ’	none
t	**426**	49		1	1				3
t^{h}	6	**470**	2	1					1
t’		2	**476**		2				
s		1		**475**	1				3
s’		2		1	**474**			2	1

b. CVs spliced from PP-initial position

	t	t^{h}	t’	s	s’	ʧ	ʧ^{h}	ʧ’	none
t	**462**	16	1						1
t^{h}	10	**466**	3	1					
t’	6	1	**471**						2
s		1		**473**	3		2		1
s’	1		1	1	**476**	1			

c. CVs spliced from Wd-initial position

	t	tʰ	t'	s	s'	ʧ	ʧʰ	ʧ'	none
t	**386**	5	5	1		57			26
tʰ	37	**412**	1	5		1	20		4
t'	5	2	**467**						6
s	17	7		**428**		16			12
s'	1				**479**				

The results of confusion matrices revealed that the overall patterns of confusions were different across prosodic domains where the target CVs were extracted from. When lenis stop CVs were spliced from IP-initial position, there was more confusion with aspirated stops. In the identification of CVs taken from PP-initial position, there was relatively small confusion between lenis and aspirated stops. When the CVs were spliced from Wd-initial positions, lenis stops were confused with lenis affricate /ʧ/ or none. In the identification of aspirated stops from Wd-initial position, there was confusion with lenis stop /t/ and aspirated affricate / ʧʰ/. It has been already noted that VOT values of lenis stops in IP and PP-initial position overlapped with those of aspirated stops in Wd-initial positions. The greater durational variation of lenis stops caused more confusion with aspirated stops in the higher prosodic domains. For aspirated stops, the shorter duration in Wd-initial position caused more confusion with lenis stops.

For lenis fricatives, listeners identified target CVs as lenis stop, lenis affricate or none in the identification of target CVs taken from Wd-initial positions. The results revealed that there was no confusion between lenis and tense fricatives across prosodic domains.

On the contrary, tense stops and fricatives did not show systematic confusion with other segments, which was reflected by the error rates. The less misidentification of tense stops and fricatives across prosodic domains can be explained by a ceiling effect. Since tense categories are already in the extreme position of hyperarticulation and therefore highly distinctive, there is no 'room' for further enhancement of acoustic properties in the position of higher information content.

In general, this study provides evidence that speakers modulate their speech clarity depending on information content. By enhancing acoustic properties and contrast, speakers tend to provide perceptual cues for the prosodic position with less contextual information although the enhancing strategies are different among phonation types. However, the enhanced properties in the prosodic domain-initial segments might also reflect the hierarchical organization of the prosodic constituents. In the previous studies of prosodic domain-initial strengthening, it has been claimed that the enhanced phonetic properties of segments in the domain-initial positions could possibly tell the listener about the strength of the prosodic boundary, similar to the way that listeners could use degree of final lengthening at the boundary (Fougeron (2001), Cho & Keating (2001)).

With regard to the final lengthening, it has been suggested that consonants or vowels in phrase-final positions are longer in duration than those located phrase medially (Keating et al. (1999), Byrd et al. (2005)). The final lengthening was considered to mark the edge of the prosodic boundary. In the current study, final vowel was lengthened cumulatively when the position moves up in the hierarchy and the final vowel lengthening showed more robust distinction across prosodic domains. It did not vary with phonation types for stops and fricatives following the target vowels, which was contrary to the prosodic domain-initial properties of stops and fricatives.

It is not yet determined whether the prosodically conditioned properties in the prosodic domain-initial segments provide perceptual cues for higher information content or for prosodic boundary or both of them. So, it is necessary to further investigate the role of the enhanced properties of prosodic domain-initial segments in speech processing.

4.3 Remaining Questions

There was inconsistency across individual subjects in revealing acoustic distinctions as a function of prosodic domain boundary, and the three prosodic domains (IP, PP and Wd) were not consistently distinguished by acoustic properties. It has already been shown that the same distinction of prosodic constituents was not made by all speakers or for all segments (Fougeron & Keating (1997), Cho & Keating (2001)). In the current study, the distinction of prosodic domains was variable in prosodic domain-initial positions as well as prosodic domain-final positions. Since prosodic phrasing plays a role in information processing by both speakers and listeners, suprasegmental properties seem to provide salient cues about the boundary of prosodic domains in addition to the temporal properties of segments. However, it is not yet clear how the variability across speakers and the distinction of prosodic domains are perceived by listeners.

In addition, the less consistent word type effect on the identification of lenis fricatives seems to be caused by less variation between the two word types in speech stimuli from speaker M1. Different from the results in the pooled data, speaker M1 did not show significant word type effect on fricatives. This result seems to reflect talker variance and it implies that talker variance in the distinction of prosodic domains could affect the perception of the prosodic domain-initial segments. In order to find out how the between-subject variations affect the perception of segments across prosodic domains, a perception study with stimuli from multiple talkers is necessary.

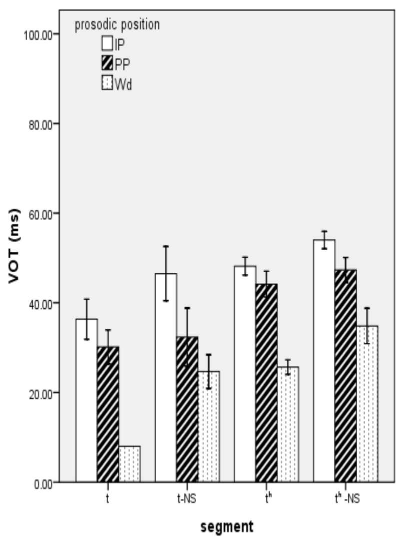
Since the enhanced prosodic domain-initial properties were found to be perceptually distinguished by listeners, it may facilitate lexical access to the prosodic domain-initial lexical item with less contextual information. However, it still remains to be unequivocally documented how Korean listeners use the prosodically conditioned properties in lexical segmentation in continuous speech.

Appendix A. Results of Individual Subjects in the Production Study

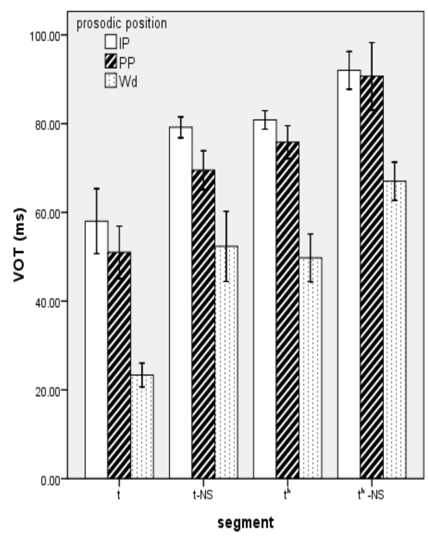
1. Stops

(1) VOT[22]

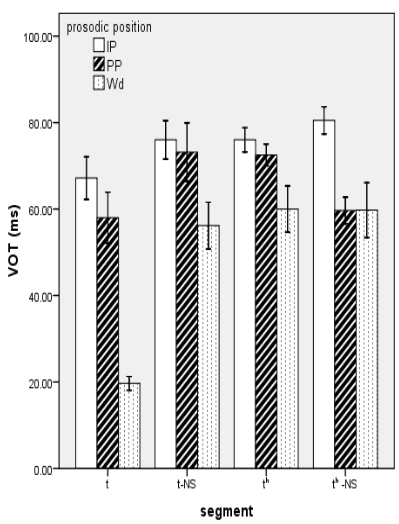
M1



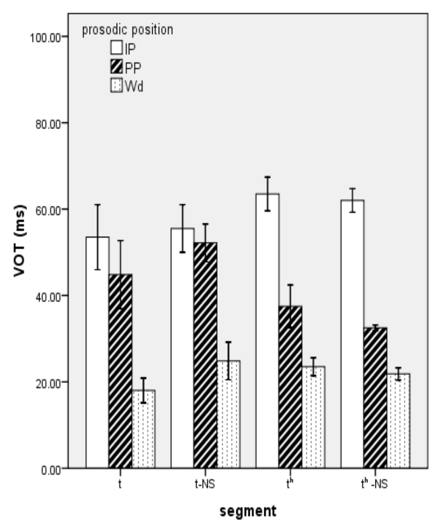
M2



M3

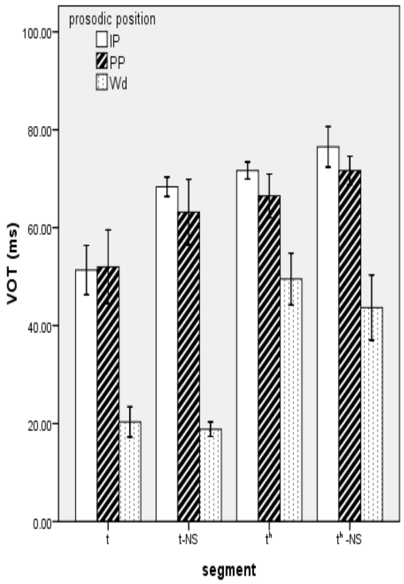


F1

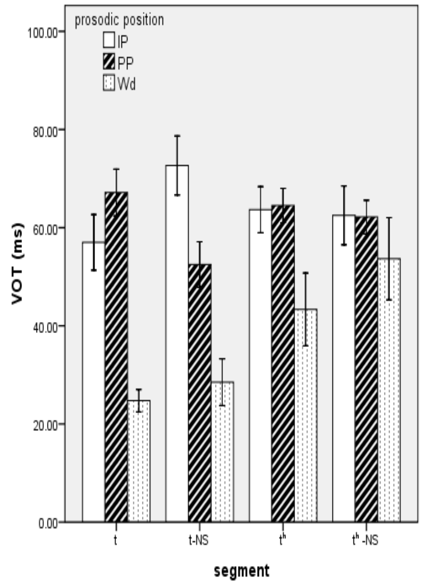


[22] The results of VOT for fortis stops /t'/ did not show consistent variation across prosodic domain-initial positions, I excluded the results in the figure.

F2

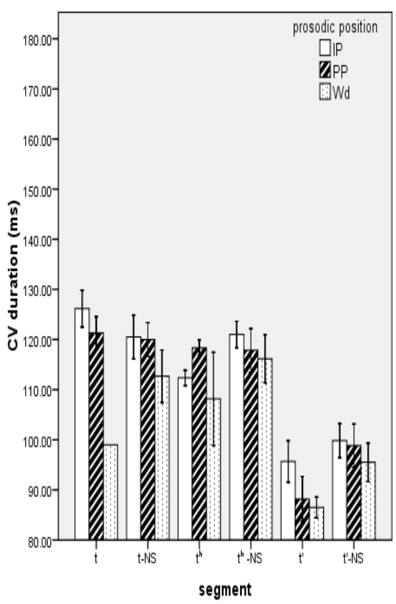


F3

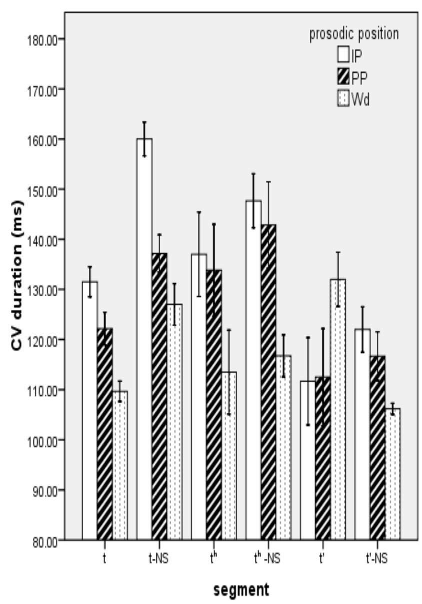


(2) CV duration

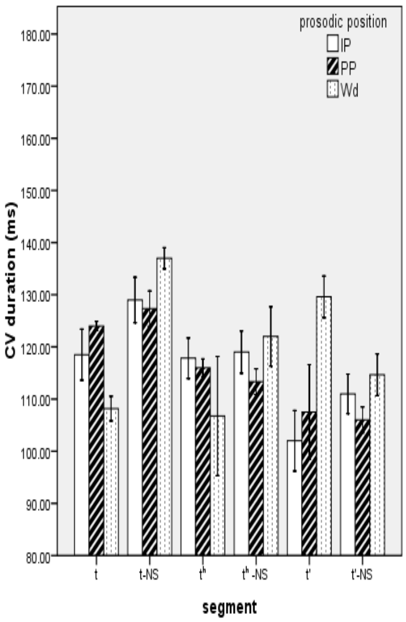
M1



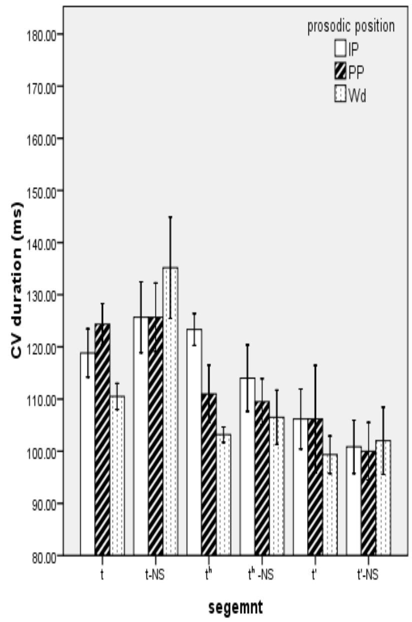
M2



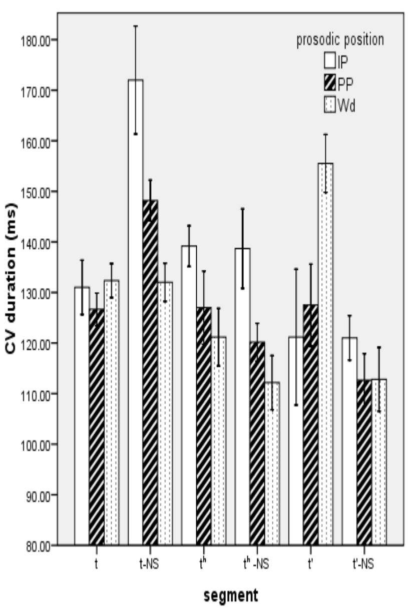
M3


prosodic position
IP
PP
Wd
CV duration (ms)
180.00
170.00
160.00
150.00
140.00
130.00
120.00
110.00
100.00
90.00
80.00
t
t-NS
tʰ
tʰ-NS
t'
t'-NS
segment

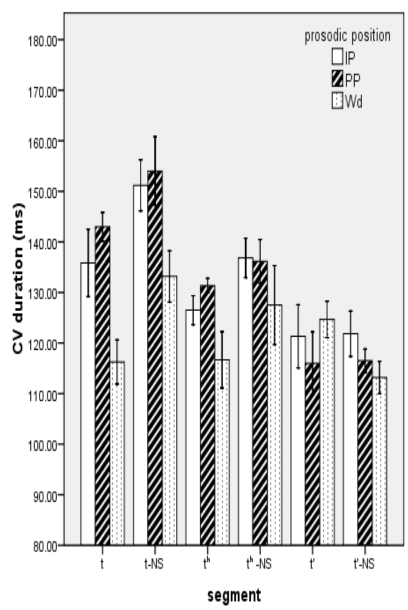
F1


prosodic position
IP
PP
Wd
CV duration (ms)
180.00
170.00
160.00
150.00
140.00
130.00
120.00
110.00
100.00
90.00
80.00
t
t-NS
tʰ
tʰ-NS
t'
t'-NS
segemnt

F2

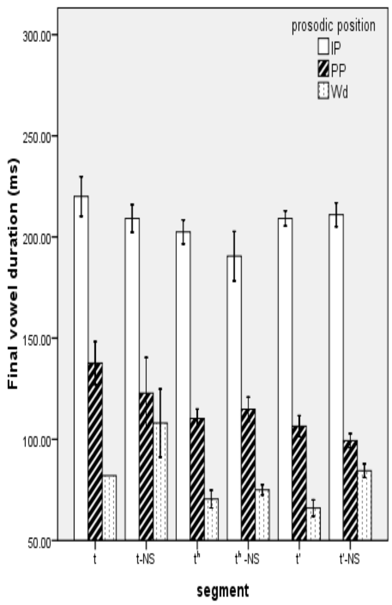

prosodic position
IP
PP
Wd
CV duration (ms)
180.00
170.00
160.00
150.00
140.00
130.00
120.00
110.00
100.00
90.00
80.00
t
t-NS
tʰ
tʰ-NS
t'
t'-NS
segment

F3

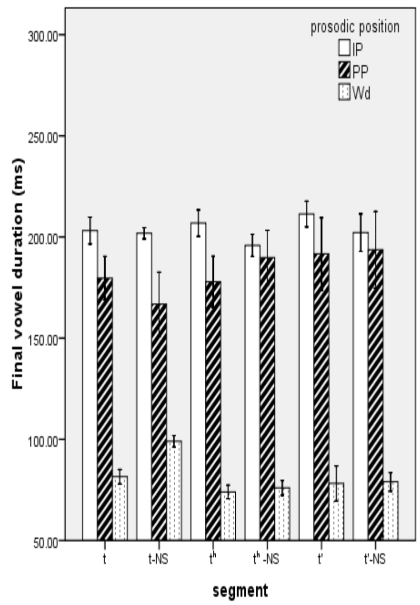

prosodic position
IP
PP
Wd
CV duration (ms)
180.00
170.00
160.00
150.00
140.00
130.00
120.00
110.00
100.00
90.00
80.00
t
t-NS
tʰ
tʰ-NS
t'
t'-NS
segment

(3) Final vowel duration

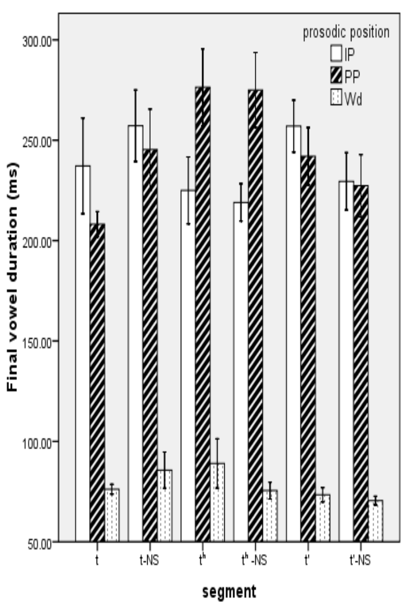
M1



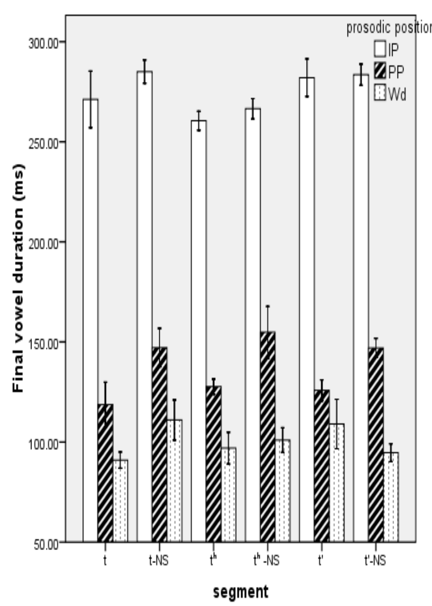
M2



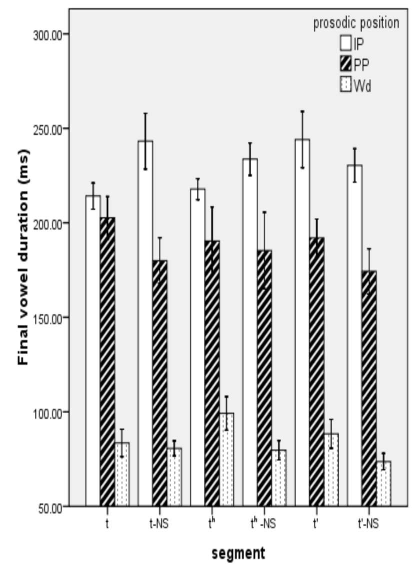
M3



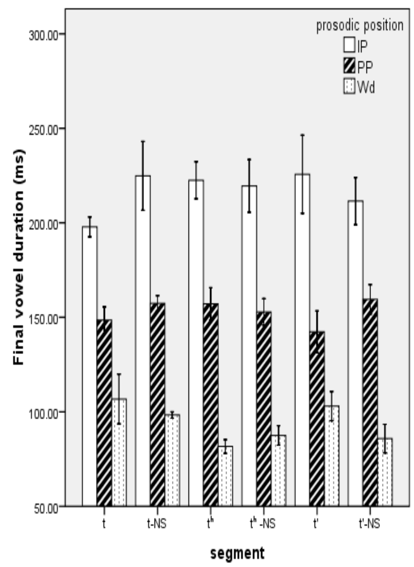
F1



F2

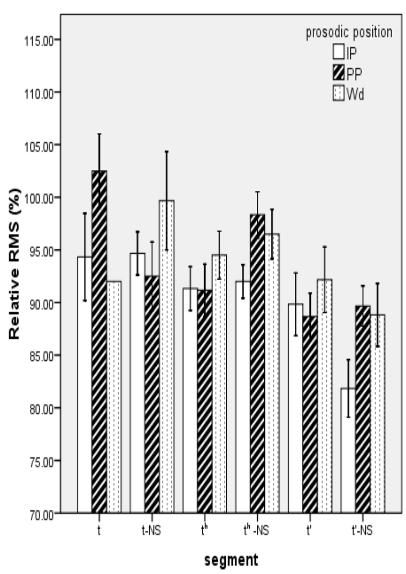


F3

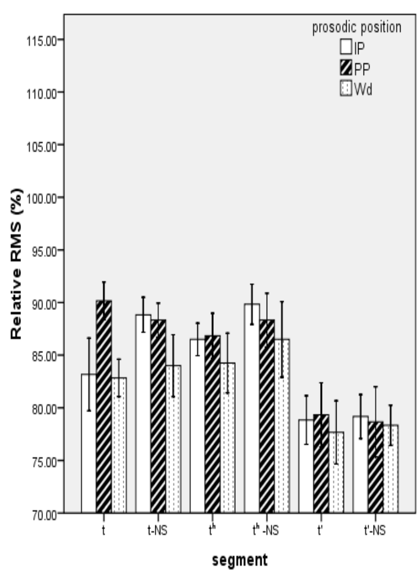


(4) Relative RMS burst energy

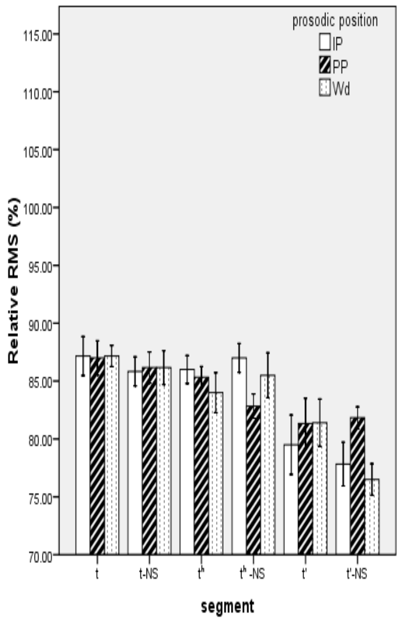
M1



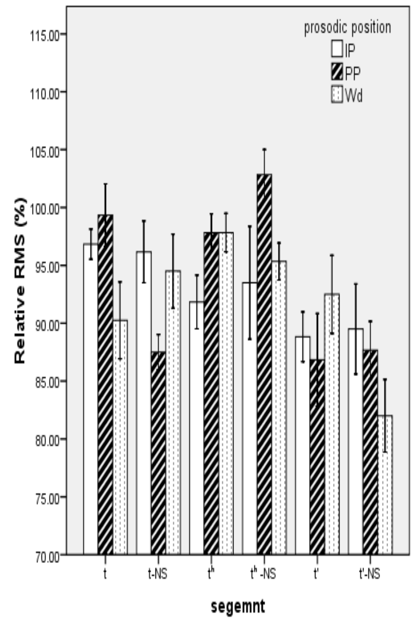
M2



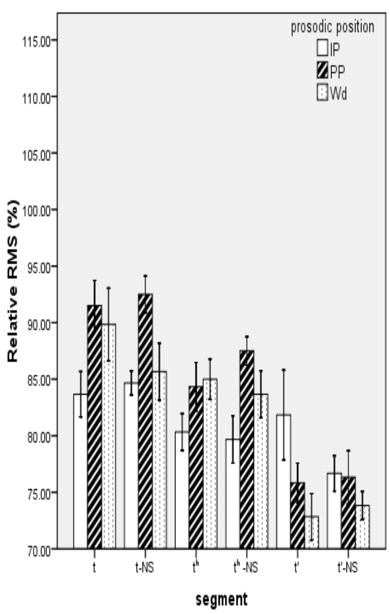
M3


prosodic position
IP
PP
Wd
Relative RMS (%)
115.00
110.00
105.00
100.00
95.00
90.00
85.00
80.00
75.00
70.00
t
t-NS
tʰ
tʰ -NS
t'
t'-NS
segment

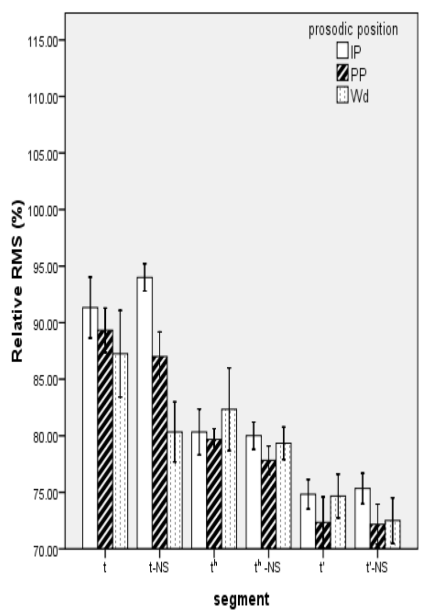
F1


prosodic position
IP
PP
Wd
Relative RMS (%)
115.00
110.00
105.00
100.00
95.00
90.00
85.00
80.00
75.00
70.00
t
t-NS
tʰ
tʰ -NS
t'
t'-NS
segemnt

F2

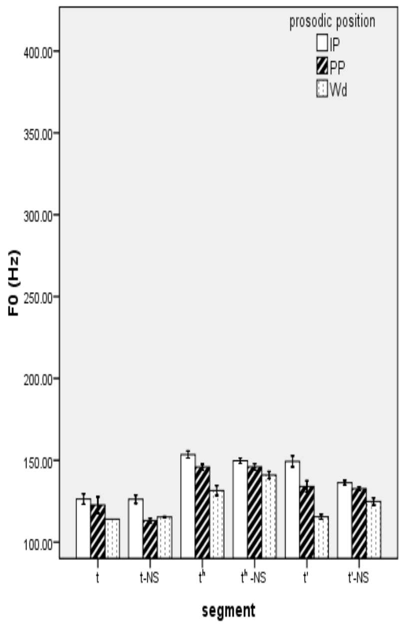

prosodic position
IP
PP
Wd
Relative RMS (%)
115.00
110.00
105.00
100.00
95.00
90.00
85.00
80.00
75.00
70.00
t
t-NS
tʰ
tʰ -NS
t'
t'-NS
segment

F3

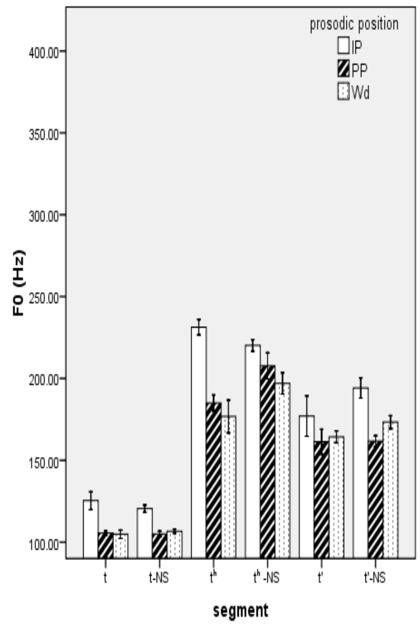

prosodic position
IP
PP
Wd
Relative RMS (%)
115.00
110.00
105.00
100.00
95.00
90.00
85.00
80.00
75.00
70.00
t
t-NS
tʰ
tʰ -NS
t'
t'-NS
segment

(5) F0

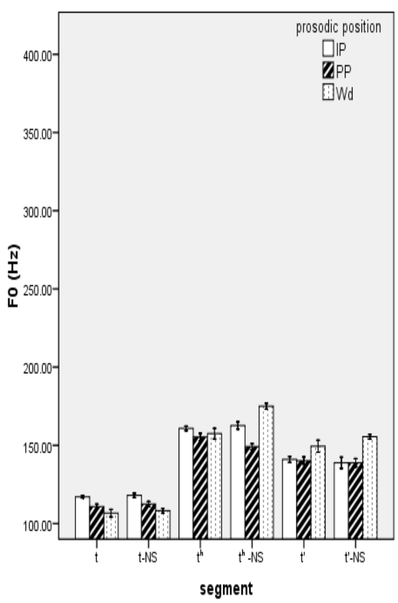
M1



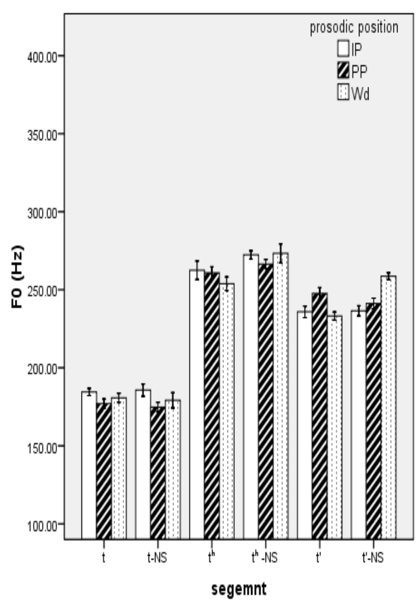
M2



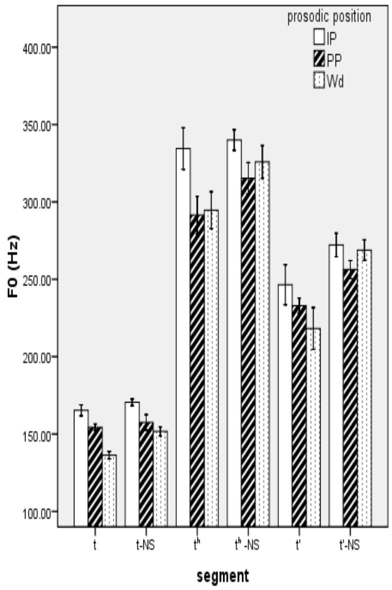
M3



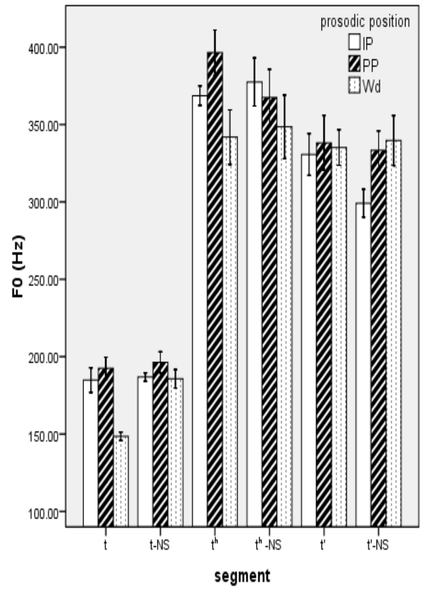
F1



F2

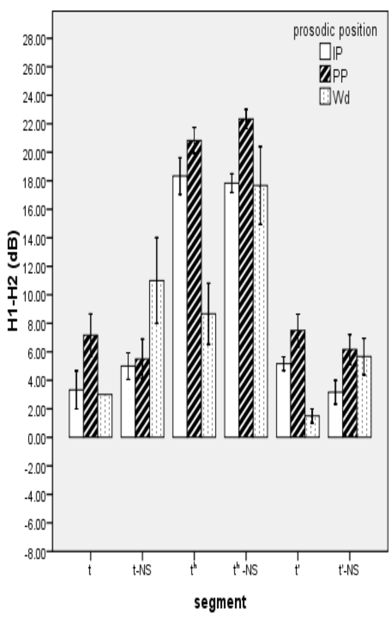


F3

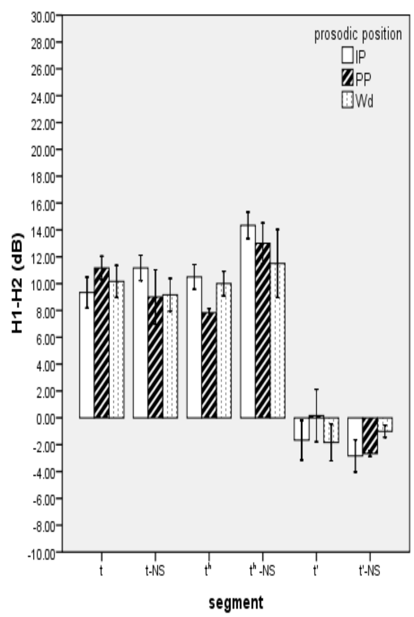


(6) H1-H2

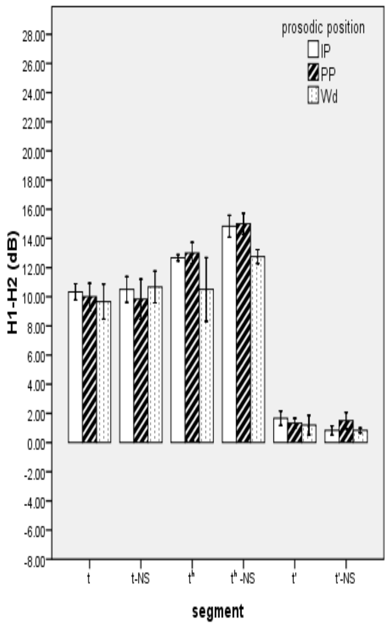
M1



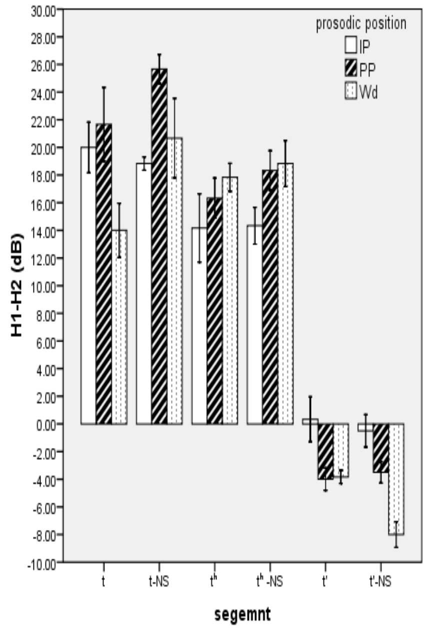
M2



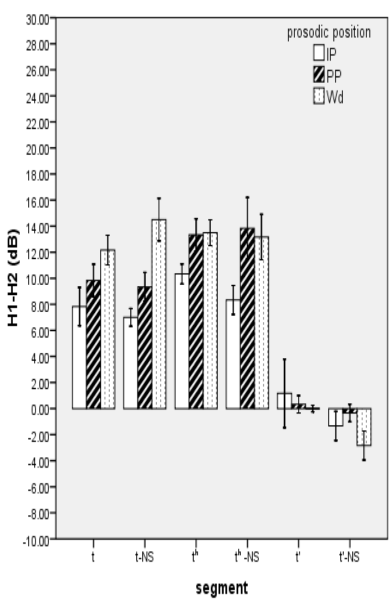
M3


prosodic position
IP
PP
Wd
H1-H2 (dB)
t
t-NS
tʰ
tʰ-NS
t'
t'-NS
segment

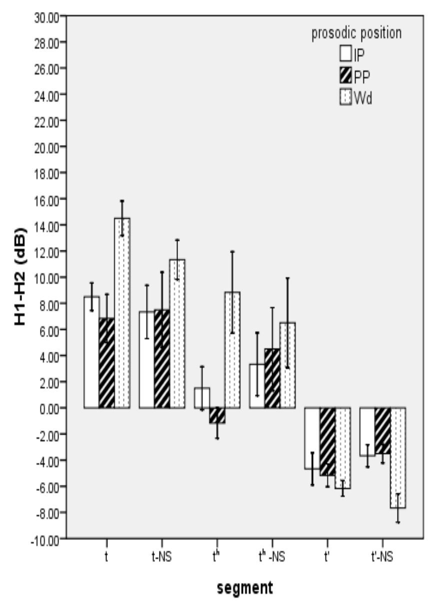
F1


prosodic position
IP
PP
Wd
H1-H2 (dB)
t
t-NS
tʰ
tʰ-NS
t'
t'-NS
segemnt

F2


prosodic position
IP
PP
Wd
H1-H2 (dB)
t
t-NS
tʰ
tʰ-NS
t'
t'-NS
segment

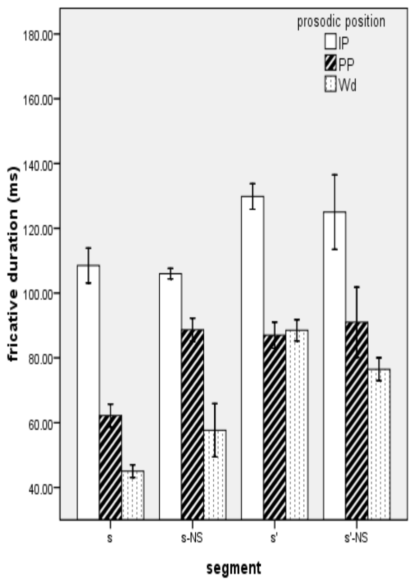
F3


prosodic position
IP
PP
Wd
H1-H2 (dB)
t
t-NS
tʰ
tʰ-NS
t'
t'-NS
segment

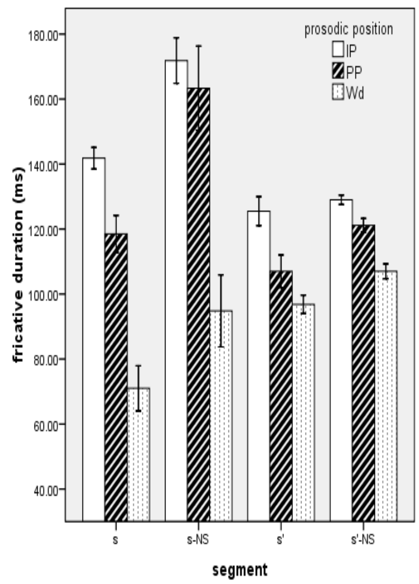
2. Fricatives

(1) Fricative duration

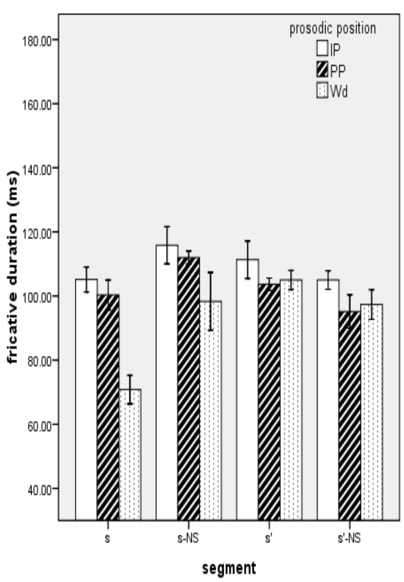
M1



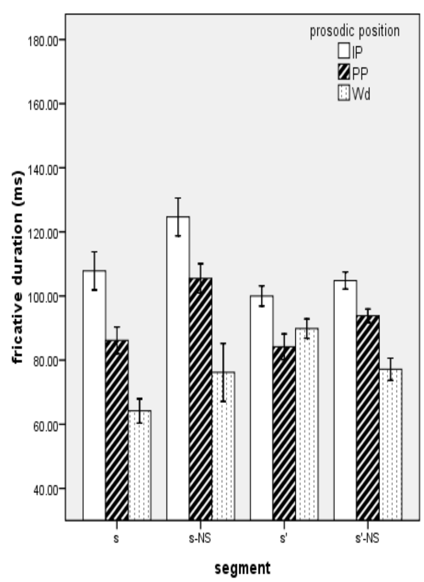
M2



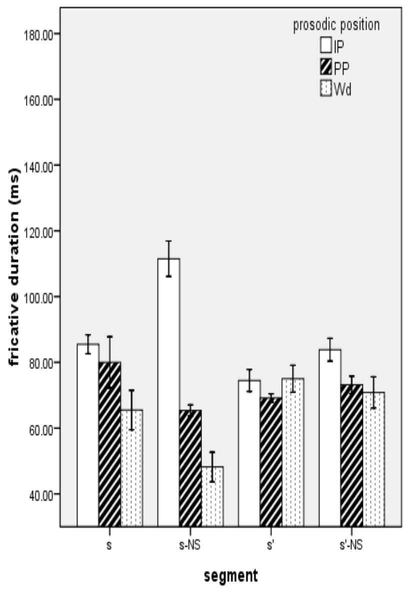
M3



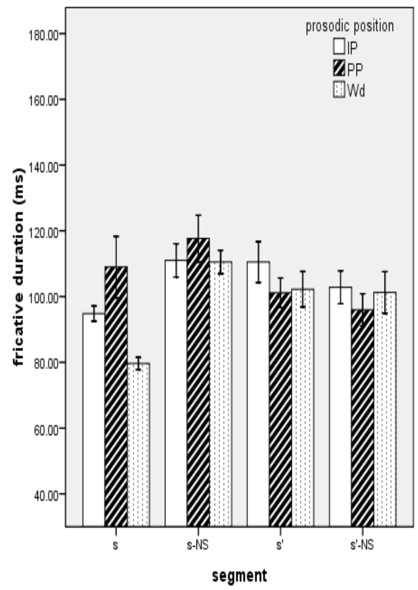
F1



F2

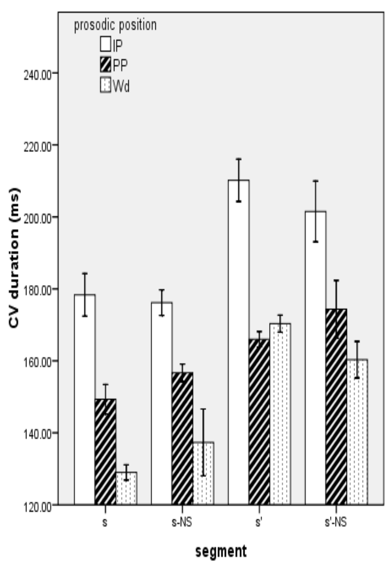


F3

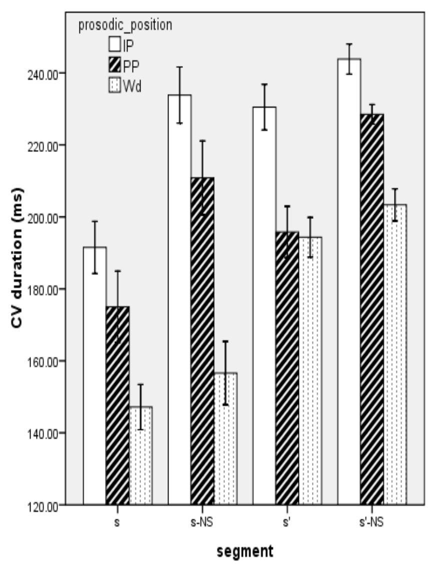


(2) CV duration

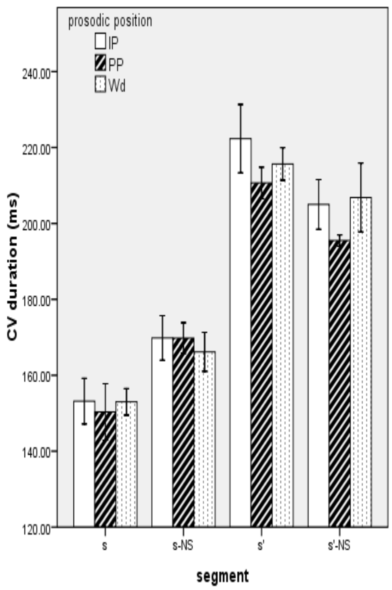
M1



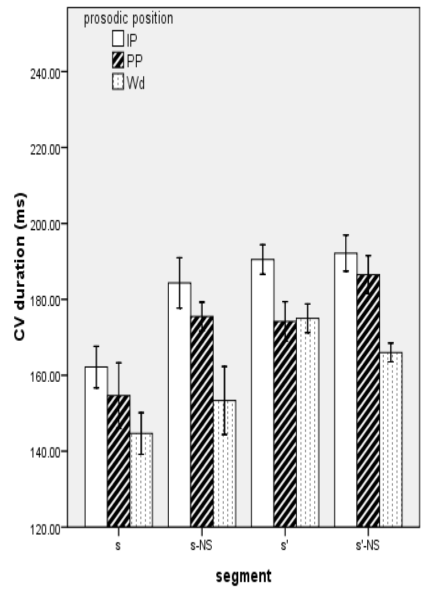
M2



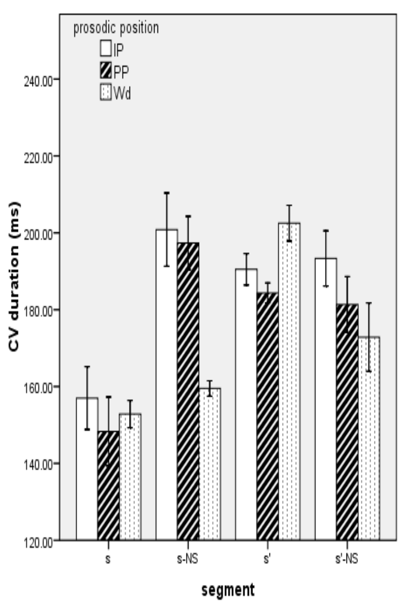
M3


prosodic position
IP
PP
Wd
240.00
220.00
200.00
180.00
160.00
140.00
120.00
CV duration (ms)
s
s-NS
s'
s'-NS
segment

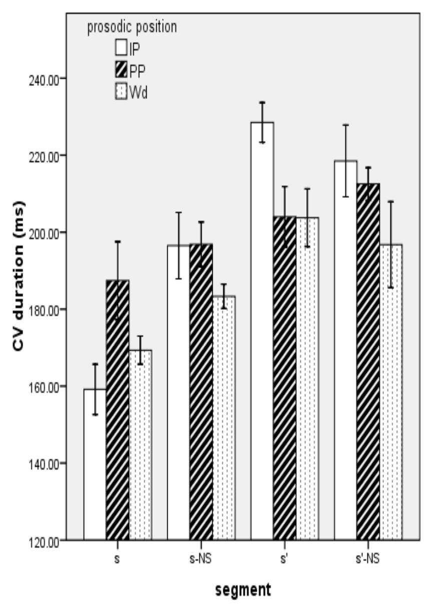
F1


prosodic position
IP
PP
Wd
240.00
220.00
200.00
180.00
160.00
140.00
120.00
CV duration (ms)
s
s-NS
s'
s'-NS
segment

F2

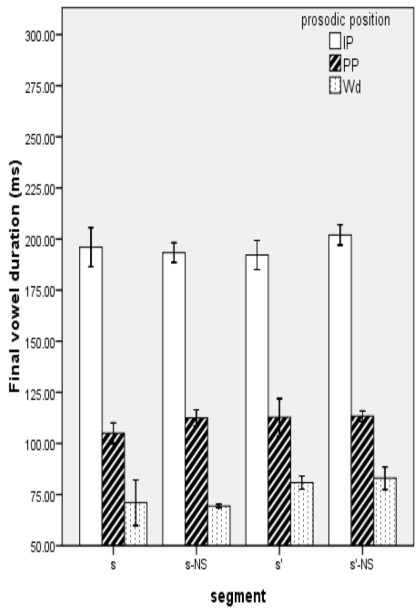

prosodic position
IP
PP
Wd
240.00
220.00
200.00
180.00
160.00
140.00
120.00
CV duration (ms)
s
s-NS
s'
s'-NS
segment

F3

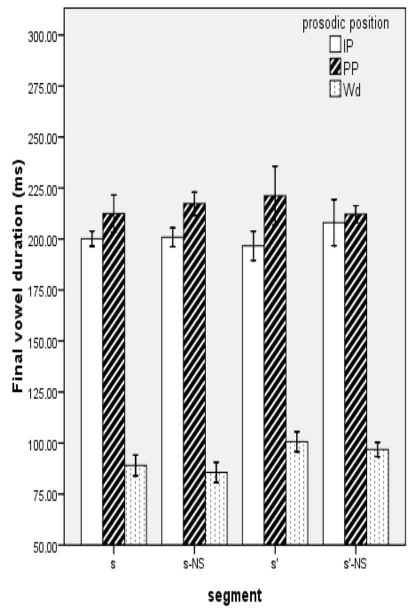

prosodic position
IP
PP
Wd
240.00
220.00
200.00
180.00
160.00
140.00
120.00
CV duration (ms)
s
s-NS
s'
s'-NS
segment

(3) Final vowel duration

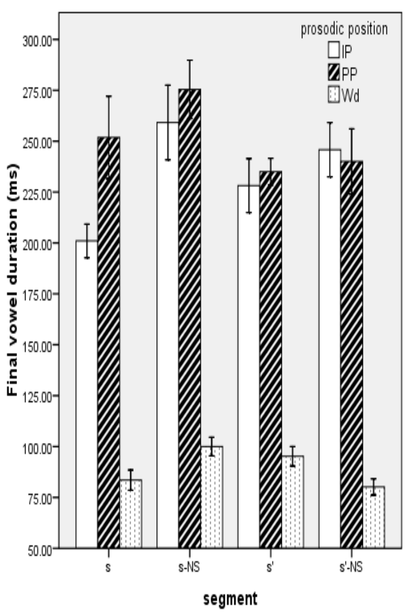
M1



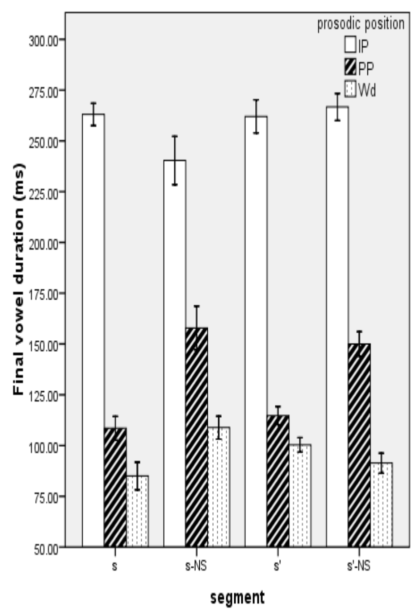
M2



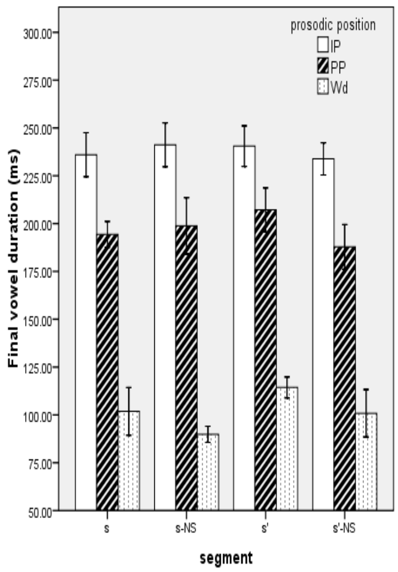
M3



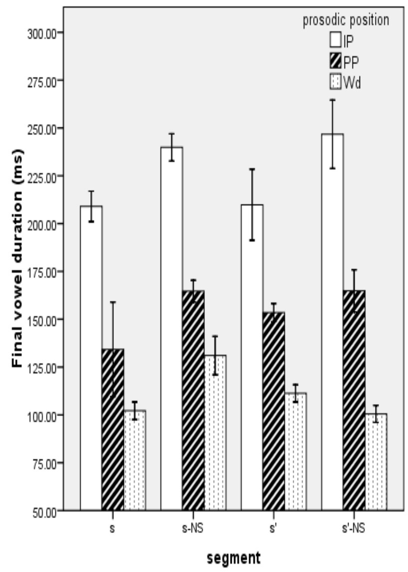
F1



F2

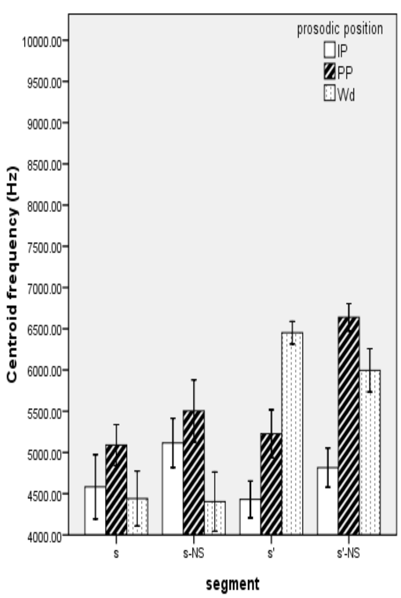


F3

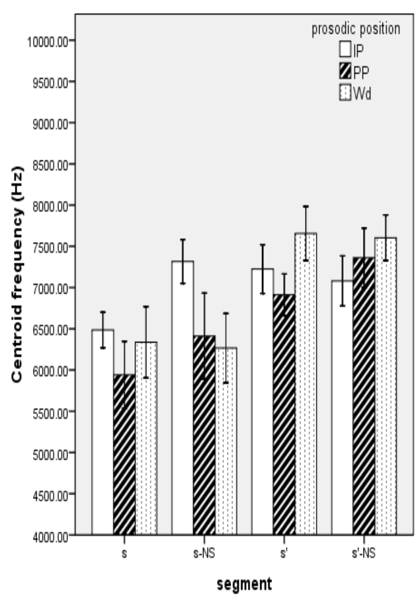


(4) Centroid frequency

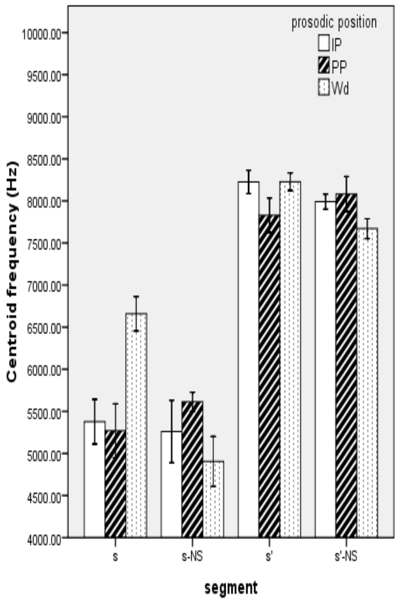
M1



M2

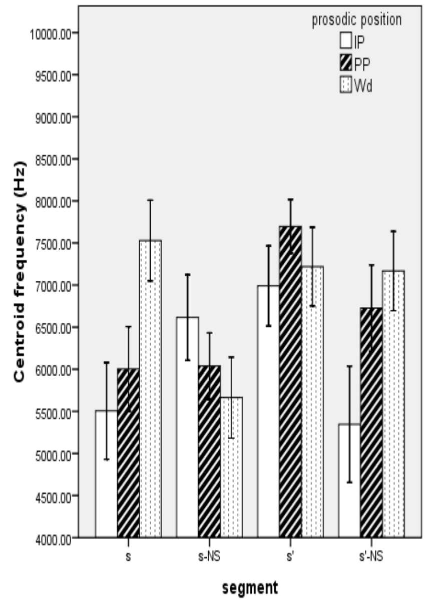


M3



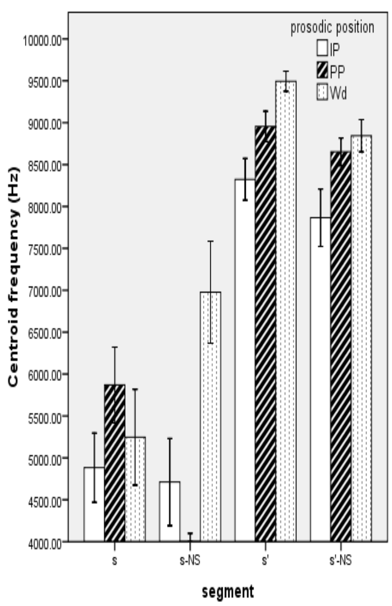
prosodic position
IP
PP
Wd
Centroid frequency (Hz)
segment
s
s-NS
s'
s'-NS

F1



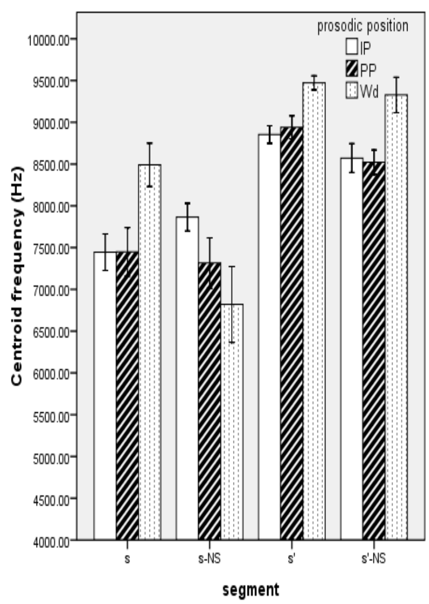
prosodic position
IP
PP
Wd
Centroid frequency (Hz)
segment
s
s-NS
s'
s'-NS

F2



prosodic position
IP
PP
Wd
Centroid frequency (Hz)
segment
s
s-NS
s'
s'-NS

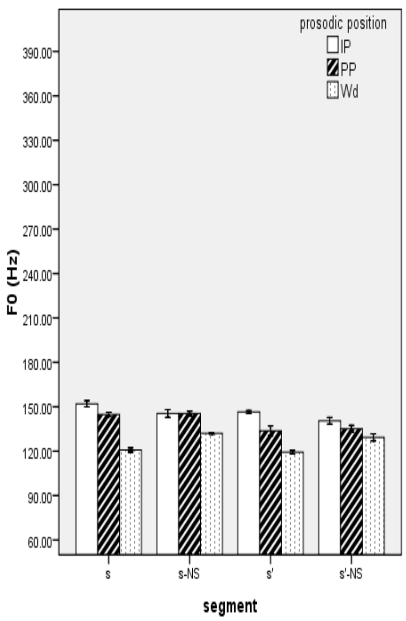
F3



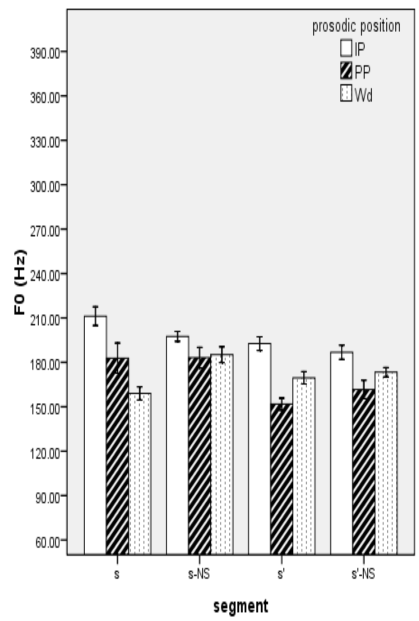
prosodic position
IP
PP
Wd
Centroid frequency (Hz)
segment
s
s-NS
s'
s'-NS

(5) F0

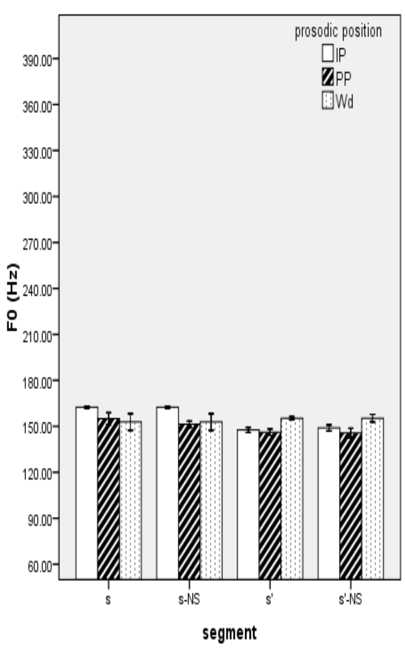
M1



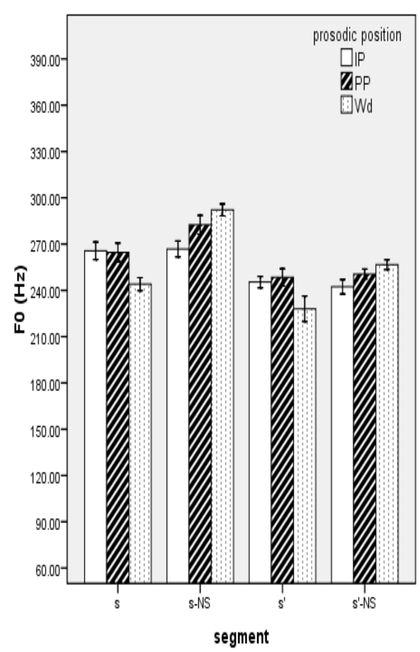
M2



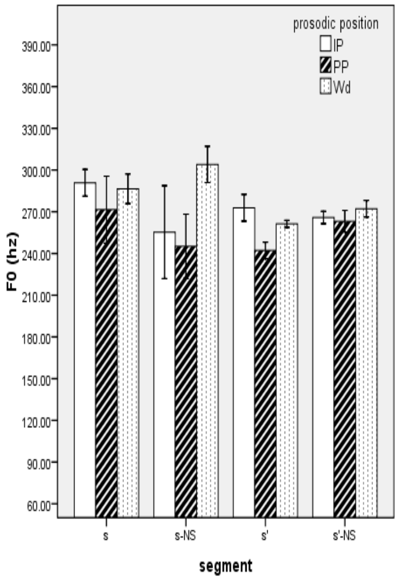
M3



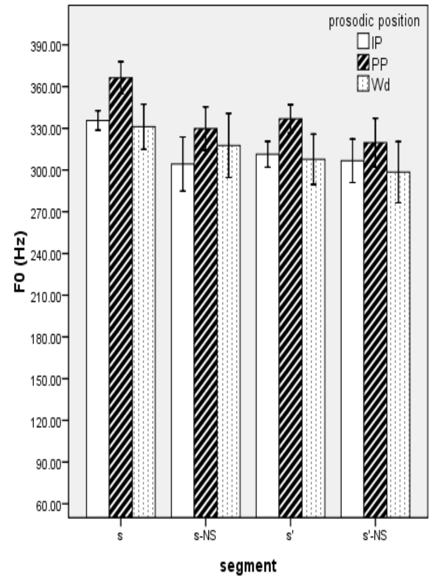
F1



F2

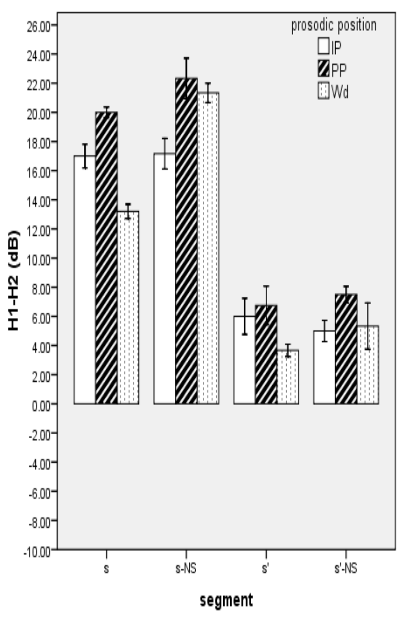


F3

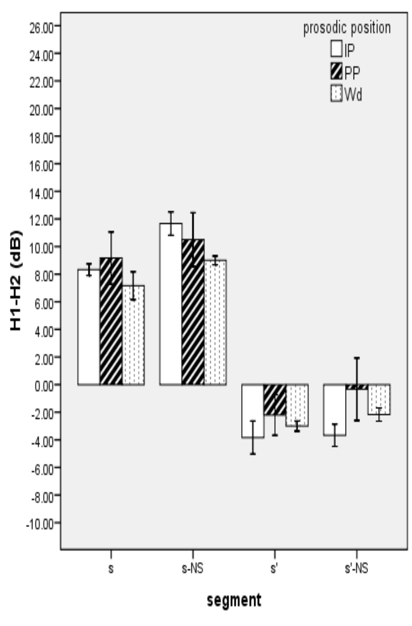


(6) H1-H2

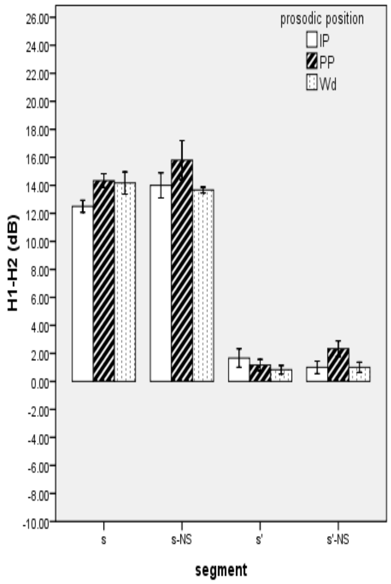
M1



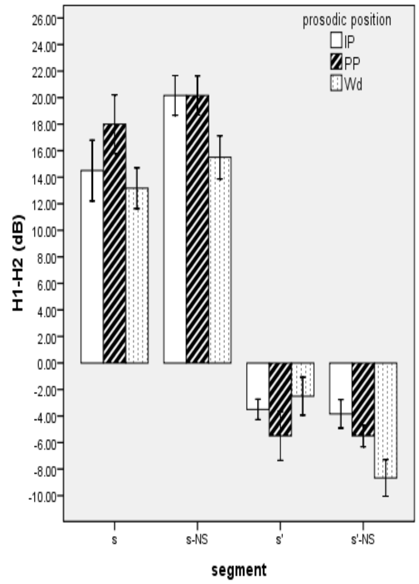
M2



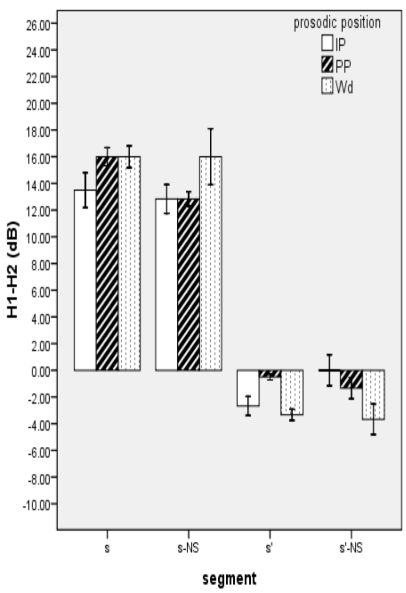
M3


prosodic position
IP
PP
Wd
H1-H2 (dB)
s
s-NS
s'
s'-NS
segment

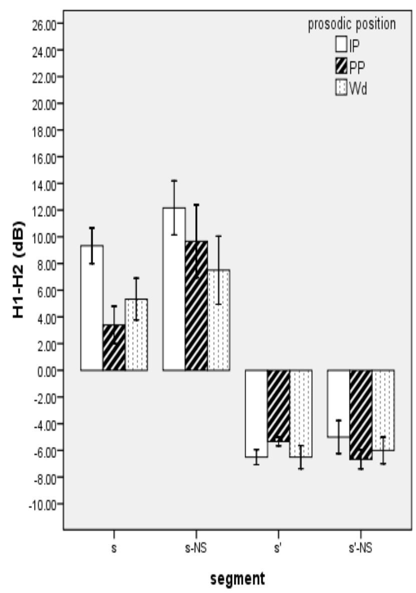
F1


prosodic position
IP
PP
Wd
H1-H2 (dB)
s
s-NS
s'
s'-NS
segment

F2


prosodic position
IP
PP
Wd
H1-H2 (dB)
s
s-NS
s'
s'-NS
segment

F3


prosodic position
IP
PP
Wd
H1-H2 (dB)
s
s-NS
s'
s'-NS
segment

Appendix B. Stimuli

1. Test sentences for /t, t^h, t'/

1.1 Real words

(1) IP-initial

(a) /t/

(i) [IP Nae doŋseŋdɯ-ɾa], [IP tat^{h}ɯ-lɯl təndʒjəbwa-ɾa]

My brothers-VOC, dart-Acc throw-IMP

'My brothers, throw the dart'

(ii) [IP Nae doŋseŋdɯ-ɾa], [IP tasugjəre t'aɾa-ɾa]

My brothers-VOC, decision by majority follow-IMP

'My brothers, accept a majority decision'

(iii) [IP Nae doŋseŋdɯ-ɾa], [IP tasugjəllo kjəltʃ'əŋ-he]

My brothers-VOC, decision by majority decide-IMP

'My brothers, decide by majority'

(b) /t^h/

(i) [IP Nae doŋseŋdɯ-ɾa], [IP t^{h}adʒo-lɯl t'aɾaga-bwa]

My brothers-VOC, ostrich-Acc follow-IMP

'My brothers, follow the ostrich'

(ii) [IP Nae doŋseŋdɯ-ɾa], [IP t^{h}adʒan noɾi-lɯl he-ɾa]

My brothers-VOC, Tarzan play-Acc do-IMP

'My brothers, play Tarzan'

(iii) [IP Nae doŋseŋdɯ-ɾa], [IP t^{h}adʒoŋ-ɯl poɾəga-ɾa]

My brothers-VOC, striking of the bell see-go-IMP

'My brothers, go to see the striking of the bell'

(c) /t'/

(i) [IP Nae doŋseŋdɯ-ɾa], [IP t'at'ɯthe ʧigois'-ə]

My brothers-VOC, warm become-DEC

'My brothers, it is getting warmer'

(ii) [IP Nae doŋseŋdɯ-ɾa], [IP t'aʤiʤi ʧom maɾa-ɾa]

My brothers-VOC, inquire into a doubtful point NEG-IMP

'My brothers, don't inquire into a doubtful point'

(iii) [IP Nae doŋseŋdɯ-ɾa], [IP t'asɯhage ibə-ɾa]

My brothers-VOC, warm take on-IMP

'My brothers, take your warm clothes'

(2) PP-initial

(a) /t/

(i) [IP [PP Sosimhan minsu-ka] [PP tasi toʤən-ɯl he-s'ə]]

Timid Minsu-Nom again challenge-Acc do-Past

'A timid Minsu challenged again'

(ii) [IP [PP həjakhan minsu-ka] [PP tasima-lɯl məgə-jo]]

weak Minsu-Nom kelp-Acc eat-DEC

'A weak Minsu eats kelp'

(iii) [IP [PP Sosimhan minsu-ka] [PP tasugjəl-ɯl paɾe-jo]]

Timid Minsu-Nom decision by majority-Acc hope-DEC

'A timid Minsu wants to decide by majority'

(b) /t^h/

(i) [IP [PP səŋsilhan minsu-ka] [PP t^haʤak-ɯl siʤak-he-s'ə]] diligent Minsu-Nom harvest-Acc start-do-Past ending

'A diligent Minsu starts his harvest'

(ii) [IP [PP Sosimhan minsu-ka] [PP tʰaʤo-lɯl siɾəhe-jo]]

Timid Minsu-Nom ostrich-Acc hate-DEC

'A timid Minsu hates the ostrich'

(iii) [IP [PP sinʤuŋhan minsu-ka] [PP tʰaʤo- lɯl ʧ'igə-s'ə-jo]]

cautious Minsu-Nom ostrich-Acc take a picture-DEC

'A cautious Minsu takes a picture of the ostrich'

(c) /t'/

(i) [IP [PP Sosimhan minsu-ka] [PP t'aʤiɾjə hes'ə-s'ə-jo]]

Timid Minsu-Nom distinguish between right and wrong do-past-DEC

'A timid Minsu tried to distinguish between right and wrong'

(ii) [IP [PP həjakhan minsu-ka] [PP t'at'ɯthan kəs-ɯl məg-ə]]

weak Minsu-Nom warm thing-Acc eat-DEC

'A weak Minsu eats warm things'

(iii) [IP [PP Sosimhan minsu-ka] [PP t'agwi-lɯl maʤa-s'ə-jo]]

Timid Minsu-Nom a cheek-Acc be boxed-past-DEC

'A timid Minsu got boxed on the ear'

(3) Wd-initial

(a) /t/

(i) [IP minjəŋi-ka [PP ap'a [Wd tatʰɯ-lɯl] pusjə-s'ə-jo]]

minyoung-Nom father ('s) dart-Acc break-past-DEC

'Miyoung broke her father's dart'

(ii) [IP minjəŋi-ka [PP ap'a [Wd tat^{h}ɯ-lɯl] pərə-s'ə-jo]]

minyoung-Nom father ('s) dart-Acc throw -past-DEC

'Miyoung threw her father's dart'

(iii) [IP minjəŋi-ka [PP op'a [Wd tat^{h}ɯ-lɯl] pərə-s'ə-jo]]

minyoung-Nom brother('s) dart-Acc throw -past-DEC

'Miyoung threw her brother's dart'

(b) /t^h/

(i) [IP minjəŋi-ka [PP ap'a [Wd t^{h}aʤo-lɯl] ʧ'igə-s'ə-jo]]

Minyoung-Nom father ostrich-Acc take a picture-past-DEC

'Minyoung took a picture of a daddy ostrich'

(ii) [IP minjəŋi-ka [PP ima [Wd t^{h}abaksaŋ-ɯl] ibə-s'ə]]

Minyoung-Nom forehead bruise-Acc get-past-DEC

Minyoung got a bruise on her forehead'

(iii) [IP minjəŋiga [PP əmma [Wd t^{h}aʤo-lɯl] ʧ'igə-s'ə-jo]]

Minyoung-Nom mother ostrich-Acc take a picture-past-DEC

'Minyoung took a picture of a mommy ostrich'

(c) /t'/

(i) [IP minjəŋi-ka [PP pak'a [Wd t'agwi- lɯl] t'eɾjə-s'ə-jo]]

Minyoung-Nom Park cheek-Acc slap-past-DEC

'Minyoung slapped Park on the cheek'

(ii) [IP minjəŋiga [PP kimga [Wd t'agwi- lɯl] t'eɾjə-s'ə-jo]]

Minyoung-Nom Kim cheek-Acc slap-past-DEC

'Minyoung slapped Kim on the cheek'

(iii) [IP minjəŋiga [PP op'a [Wd t'agwi-lɯl] t'eɾjə-s'ə-jo]]

Minyoung-Nom brother cheek-Acc slap-past-DEC

'Minyoung slapped her brother on the cheek'

1.2. Nonsense words

(1) IP-initial

(a) /t/

(i) [IP Nae doŋseŋdɯ-ɾa], [IP tamale-ɾa mal-he-bwa]

My brothers-VOC, Tamal-as say-try-DEC

'My brothers, try to say 'Tamal''

(ii) [IP Nae doŋseŋdɯ-ɾa], [IP tasuni-ɾa mal-he-bwa]

My brothers-VOC, Tasun-as say-try-DEC

'My brothers, try to say 'Tasun''

(iii) [IP Nae doŋseŋdɯ-ɾa], [IP taʤuni-ɾa mal-he-bwa]

My brothers-VOC, Tazun-as say-try-DEC

'My brothers, try to say 'Tazun''

(b) /t^h/

(i) [IP Nae doŋseŋdɯ-ɾa], [IP t^hamali-ɾa mal-he-bwa]

My brothers-VOC, Thamal-as say-try-DEC

'My brothers, try to say 'Thamal''

(ii) [IP Nae doŋseŋdɯ-ɾa], [IP t^hasuni-ɾa mal-he-bwa]

My brothers-VOC, Thasun-as say-try-DEC

'My brothers, try to say 'Thasun''

(iii) [IP Nae doŋseŋdɯ-ɾa], [IP tʰaʤuni-ɾa mal-he-bwa]

My brothers-VOC, Thazun-as say-try-DEC

‘My brothers, try to say ‘Thazun’’

(c) /t’/

(i) [IP Nae doŋseŋdɯ-ɾa], [IP t’amali-ɾa mal-he-bwa]

My brothers-VOC, Ttamal-as say-try-DEC

‘My brothers, try to say ‘Ttamal’’

(ii) [IP Nae doŋseŋdɯ-ɾa], [IP t’asuni-ɾa mal-he-bwa]

My brothers VOC, Ttasun-as say-try-DEC

‘My brothers, try to say ‘Ttasun’’

(iii) [IP Nae doŋseŋdɯ-ɾa], [IP t’aʤuni-ɾa mal-he-bwa]

My brothers- VOC, Ttazun-as say-try-DEC

‘My brothers, try to say ‘Ttazun’’

(2) PP-initial

(a) /t/

(i) [IP [PP ənulhan minsu-ka] [PP tamaliɾa ilgə-jo]]

timid Minsu-Nom Tamal-as read-DEC

‘A timid Minsu reads it as ‘Tamal’’

(ii) [IP [PP ənulhan minsu-ka] [PP tasuniɾa ilgə-jo]]

timid Minsu-Nom Tasun-as read-DEC

‘A timid Minsu reads it as ‘Tasun’’

(iii) [IP [PP ənulhan minsu-ka] [PP taʤuniɾa ilgə-jo]] timid Minsu-Nom Tazun-as read-DEC

‘A timid Minsu reads it as ‘Tazun’’

(b) /t^h/

(i) [IP [PP ənulhan minsu-ka] [PP t^hamaliɾa ilgə-jo]]

timid Minsu-Nom Thamal-as read-DEC

‘A timid Minsu reads it as ‘Thamal’’

(ii) [IP [PP ənulhan minsu-ka] [PP t^hasuniɾa ilgə-jo]]

timid Minsu-Nom Thasun-as read-DEC

‘A timid Minsu reads it as ‘Tasun’’

(iii) [IP [PP ənulhan minsu-ka] [PP t^haʤuniɾa ilgə-jo]]

timid Minsu-Nom Thazun-as read-DEC

‘A timid Minsu reads it as ‘Thazun’’

(c) /t’/

(i) [IP [PP ənulhan minsu-ka] [PP t’amaliɾa ilgə-jo]]

timid Minsu-Nom Ttamal-as read-DEC

‘A timid Minsu reads it as ‘Ttamal’’

(ii) [IP [PP ənulhan minsu-ka] [PP t’asuniɾa ilgə-jo]]

timid Minsu-Nom Ttasun-as read-DEC

‘A timid Minsu reads it as ‘Ttasun’’

(iii) [IP [PP ənulhan minsu-ka] [PP t’aʤuniɾa ilgə-jo]]

timid Minsu-Nom Ttazun-as read-DEC

‘A timid Minsu reads it as ‘Ttazun’’

(3) Wd-initial

(a) /t/

(i) [IP minjəŋi-ka [PP səŋga [Wd tamal-ɯl] pullə-jo]]

Minyoung-Nom a sacred song Tamal-Acc sing-DEC

'Minyoung is singing a sacred song, 'Tamal''

(ii) [IP minjəŋi-ka [PP səŋga [Wd tasun-ɯl] pullə-jo]]

Minyoung-Nom a sacred song Tasun-Acc sing-DEC

'Minyoung is singing a sacred song, 'Tasun''

(iii) [IP minjəŋi-ka [PP səŋga [Wd taʤun-ɯl] pullə-jo]]

Minyoung-Nom a sacred song Tazun-Acc sing-DEC

'Minyoung is singing a sacred song, 'Tazun''

(b) /t^h/

(i) [IP minjəŋi-ka [PP səŋga [Wd t^hamal-ɯl] pullə-jo]]

Minyoung-Nom a sacred song Thamal-Acc sing-DEC

'Minyoung is singing a sacred song, 'Thamal''

(ii) [IP minjəŋi-ka [PP səŋga [Wd t^hasun-ɯl] pullə-jo]]

Minyoung-Nom a sacred song Thasun-Acc sing-DEC

'Minyoung is singing a sacred song, 'Thasun''

(iii) [IP minjəŋi-ka [PP səŋga [Wd t^haʤun-ɯl] pullə-jo]]

Minyoung-Nom a sacred song Thazun-Acc sing-DEC

'Minyoung is singing a sacred song, 'Thadzun''

(c) /t'/

(i) [IP minjəŋi-ka [PP səŋga [Wd t'amal-ɯl] pullə-jo]]

Minyoung-Nom a sacred song Ttamal-Acc sing-DEC

'Minyoung is singing a sacred song, 'Ttamal''

(ii) [IP minjəŋiga [PP səŋga [Wd t'asun-ɯl] pullə-jo]]

Minyoung-Nom a sacred song Ttasun-Acc sing-DEC

'Minyoung is singing a sacred song, 'Ttasun''

(iii) [IP minjəŋiga [PP səŋga [Wd t'aʤun-ɯl] pullə-jo]]

Minyoung-Nom a sacred song Ttazun-Acc sing-DEC

'Minyoung is singing a sacred song, 'Ttazun''

2. Test sentences for /s, s'/

2.1. Real Words

(1) IP-initial

(a) /s/

(i) [IP Nae doŋseŋdɯ-ɾa], [IP saʤingi-lɯl kaʤjə-wa]

My brothers-VOC, camera-Acc bring-IMP

'My brothers, bring your camera'

(ii) [IP Nae doŋseŋdɯ-ɾa], [IP saʤin-ɯl porə-ga-ɾa]

My brothers-VOC, picture-Acc look at-go-IMP

'My brothers, go and look at the pictures'

(iii) [IP Nae doŋseŋdɯ-ɾa], [IP $satʃ^h$i-lɯl ha-ʤima-ɾa]

My brothers-VOC, luxury-Acc do-NEG-IMP

'My brothers, don't indulge in luxury'

(b) /s’/

(i) [IP Nae doŋseŋdɯ-ɾa], [IP s’age sal-sudo i’s-ə]

My brothers-VOC, cheaply buy-can be-DEC

‘My brothers, you can buy that cheaply’

(ii) [IP Nae doŋseŋdɯ-ɾa], [IP s’age-lɯl ʧunbi he-ɾa]

My brothers-VOC, blanket-Acc prepare for do-IMP

‘My brothers, prepare for a blanket’

(iii) [IP Nae doŋseŋdɯ-ɾa], [IP s’age-lɯl mandɯɾə-ɾa]

My brothers-VOC, blanket-Acc make-IMP

‘My brothers, make a blanket’

(2) PP-initial

(a) /s/

(i) [IP [PP Sosimhan minsu-ka] [PP satʃ^{h}i-ka sim-he-s’ə-jo]]

Timid minsu-Nom luxury-Nom extreme-do-past-DEC

‘A timid Minsu lived in extreme luxury’

(ii) [IP [PP Sosimhan minsu-ka] [PP sasil-ɯl mal-he-s’ə-jo]]

Timid minsu-Nom truth-Acc tell-do-past-DEC

‘A timid Minsu told the truth’

(iii) [IP [PP Sosimhan minsu-ka] [PP saʤin-ɯl pəɾjə-s’ə-jo]]

Timid minsu-Nom picture-Acc throw-past-DEC

‘A timid Minsu threw the picture’

(b) /s’/

(i) [IP [PP Sesimhan minsu-ka] [PP s’age-lɯl mandɯɾə-jo]]

Prudent Minsu-Nom blanket-Acc make-DEC

‘A prudent Minsu makes a blanket’

(ii) [IP [PP Sosimhan minsu-ka] [PP s’age-lɯl pəɾjə-s’ə-jo]]

Prudent Minsu-Nom blanket-Acc throw-past-DEC

‘A prudent Minsu threw out a blanket’

(iii) [IP [PP Sesimhan minsu-ka] [PP s’age-lɯl sawa-s’ə-jo]]

Prudent Minsu-Nom blanket-Acc buy-come-past-DEC

‘A prudent Minsu bought a blanket’

(3) Wd-initial

(a) /s/

(i) [IP minjəŋi-ka [PP pak’a [Wd saʤu-lɯl] powa-s’ə-jo]]

Minyoung-Nom Park fate-Acc see-past-DEC

‘Minyoung had Park’s fortune told by an astrologer’

(ii) [IP minjəŋi-ka [PP pak’a [Wd saʤin-ɯl] ʧ’igə-s’ə-jo]]

Minyoung-Nom Park picture-Acc take-past-DEC

‘Minyoung took Park’s picture’

(iii) [IP minjəŋi-ka [PP mina [Wd saʤin-ɯl] pəɾjə-s’ə-jo]]

Minyoung-Nom Mina picture-Acc throw-past-DEC

‘Minyoung threw Mina’s picture’

(b) /s’/

(i) [IP minjəŋi-ka [PP aga [Wd s’age-lɯl] sawa-s’ə-jo]]

Minyoung-Nom baby blanket-Acc buy-past-DEC

‘Minyoung bought the baby’s blanket’

(ii) [IP minjəŋi-ka [PP mina [Wd s’age-lɯl] pəɾjə-s’ə-jo]]

Minyoung-Nom Mina blanket-Acc throw out-past-DEC

‘Minyoung threw out Mina’s blanket’

(iii) [IP minjəŋi-ka [PP mina [Wd s’age-lɯl] madɯɾə-jo]]

Minyoung-Nom Mina blanket -Acc make-DEC

‘Minyoung makes Mina’s blanket’

2.2. Nonsense words

(1) IP-initial

(a) /s/

(i) [IP Nae doŋseŋdɯ-ɾa], [IP samali-ɾa mal-he-bwa]

My brothers-VOC, samal-as say-try-IMP

‘My brothers, try to say ‘Samal’’

(ii) [IP Nae doŋseŋdɯ-ɾa], [IP sasuni-ɾa mal-he-bwa]

My brothers-VOC, sasun-as say-try- IMP

‘My brothers, try to say ‘Sasun’’

(iii) [IP Nae doŋseŋdɯ-ɾa], [IP saʤuni-ɾa mal-he-bwa]

My brothers-VOC, sazun-as say-try- IMP

‘My brothers, try to say ‘Sazun’’

(b) /s’/

(i) [IP Nae doŋseŋdɯ-ɾa], [IP s’amali-ɾa mal-he-bwa]

My brothers-VOC, Ssamal-as say-try-IMP

‘My brothers, try to say ‘Ssamal’’

(ii) [IP Nae doŋseŋdɯ-ɾa], [IP s’asuni-ɾa mal-he-bwa]

My brothers-VOC, Ssasun-as say-try- IMP

'My brothers, try to say 'Ssasun''

(iii) [IP Nae doŋseŋdɯ-ɾa], [IP s'aʤuni-ɾa mal-he-bwa]

My brothers-VOC, Ssazun-as say-try- IMP

'My brothers, try to say Ssazun'

(2) PP-initial

(a) /s/

(i) [IP [PP ənulhan minsu-ka] [PP samali-ɾa ilgə-jo]]

timid Minsu-Nom Samal-as read-DEC

'A timid Minsu reads it as 'Samal''

(ii) [IP [PP ənulhan minsu-ka] [PP sasuni-ɾa ilgə-jo]]

timid Minsu-Nom Sasun-as read-DEC

'A timid Minsu reads it as 'Sasun''

(iii) [IP [PP ənulhan minsu-ka] [PP saʤuni-ɾa ilgə-jo]]

timid Minsu-Nom Sazun-as read-DEC

'A timid Minsu reads it as 'Sazun''

(b) /s'/

(i) [IP [PP ənulhan minsu-ka] [PP s'amali-ɾa ilgə-jo]]

timid Minsu-Nom Ssamal-as read-DEC

'A timid Minsu reads it as 'Ssamal''

(ii) [IP [PP ənulhan minsu-ka] [PP s'asuni-ɾa ilgə-jo]]

timid Minsu-Nom Ssasun-as read-DEC

'A timid Minsu reads it as 'Ssasun''

(iii) [IP [PP ənulhan minsu-ka] [PP s'aʤuni-ɾa ilgə-jo]] timid Minsu-Nom Ssazun-as read-DEC

'A timid Minsu reads it as 'Ssazun''

(3)Wd-initial

(a) /s/

(i) [IP minjəŋi-ka [PP səŋga [Wd samal-ɯl] pullə-jo]]

Minyoung-Nom a sacred song Samal-Acc sing-DEC

'Minyoung is singing a sacred song, 'Sasun''

(ii) [IP minjəŋi-ka [PP səŋga [Wd sasun-ɯl] pullə-jo]]

Minyoung-Nom a sacred song Sasun-Acc sing-DEC

'Minyoung is singing a sacred song, 'Sasun''

(iii) [IP minjəŋi-ka [PP səŋga [Wd saʤun-ɯl] pulə-jo]]

Minyoung-Nom a sacred song Sazun-Acc sing-DEC

'Minyoung is singing a sacred song, 'Sazun''

(b) /s'/

(i) [IP minjəŋi-ka [PP səŋga [Wd s'amal-ɯl] pullə-jo]]

Minyoung-Nom a sacred song Ssamal-Acc sing-DEC

'Minyoung is singing a sacred song, 'Ssamal''

(ii) [IP minjəŋi-ka [PP səŋga [Wd s'asun-ɯl] pullə-jo]]

Minyoung-Nom a sacred song Ssasun-Acc sing-DEC

'Minyoung is singing a sacred song, 'Ssasun''

(iii) [IP minjəŋi-ka [PP səŋga [Wd s'aʤun-ɯl] pullə-jo]]

Minyoung-Nom a sacred song Ssazun-Acc sing-DEC

'Minyoung is singing a sacred song, 'Ssazun''

Bibliography

Abberton, E. 1972. Some laryngeographic data for Korean stops. *Journal of International Phonetic Association*, 2, 67-78.

Abramson, A.S., and Lisker, L. 1972. "Voice-timing perception in Spanish word-initial stops," *Journal of Phonetics*, 1, 1-8.

Ahn, H. 1999. Post-Release Phonation Processes in English and Korean: Acoustic Correlates and Implications for Korean Phonology. Ph.D. dissertation. University of Texas at Austin.

Baker, K. 2002. Prosodic structure and perception of Korean domain-initial coronal stops**,** *International J. Korean Linguistics* 11, 119-132

Beckman, M.E, & Pierrehumbert, J. 1986. Intonational structure in Japanese and English, *Phonology Yearbook*, 3, 255-309.

Beckman, M.E, & Edwards, J. 1992. Intonational categories and the articulatory control of duration. In: Tohkura Y, Vatikiotis-Bateson E, Sagisaka Y, (eds.). *Speech Perception, Production, and Linguistic Structure*, 359-375. Ohmsha; Tokyo, Japan.

Bell, A. & Hooper, J. 1978. Issues and evidence in syllabic phonology. In Syllables and segments (A. Bell & J. Hooper, editors), 3-24. Amsterdam: North-Holland.

Browman, C. P., & Goldstein, L. 1992. Articulatory phonology: An overview. *Phonetica*, 49, 155-180.

Browman, C. P., & Goldstein, L. 1995. “Gestural syllable position effects in American English,” in Producing Speech: Contemporary Issues, For Katherine Harris, edited by F. Bell-Berti and L. J. Raphael (AIP Press, New York), pp. 19-33.

Byrd, D. 1994. Articualtory timing in English consonant sequences, Ph.D. thesis, University of California, Los Angeles.

Byrd, D., & Kaun, A., Narayanan, S., & Salzman, E. 2000. Phrasal signature in articulation. In M. Broe, & J. Pierrehumbert (Eds.), *Acquisition and the lexicon:*

Papers in laboratory phonology V (pp. 70-88). Cambridge, UK: Cambridge University Press.

Chang, C.-B. 2007. Korean fricatives: Production, perception, and laryngeal typology. UC Berkeley Phonology Lab Annual Report, University of California, Berkeley.

Childers, D.G. & Lee, C. K. 1991. Vowel quality factors: analysis, syntheses, and perception. *Journal of the Acoustical Society of America*, 90(5), 2394-2410.

Cho, T.-H. 2004. Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, 32, 141-176.

Cho, T.-H., & Jun, S.-A. 2000. Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean. *Chicago Linguistics Society*, 36, 31-44 (Also appeared in *UCLA Working Papers in Phonetics*, 99, 57-70).

Cho, T.-H., & Keating, P. 2001. Articulatory and acoustic studies of domain-initial strengthening in Korean. *Journal of Phonetics*, 29, 155-190.

Cho, T.-H., Jun, S., & Ladefoged, P. 2002. Acoustic and aerodynamic correlates to Korean stops and fricatives. *Journal of Phonetics*, 30(2), 193-228.

Cho, T.-H., & McQueen, J., & Cox, E. 2007. Prosodically driven phonetic detail in speech processing: The case of domain-initial strengthening in English. *Journal of Phonetics* 35(2), 210-243.

Cho, T.-H., & Keating, P. 2007. Effects of initial position versus prominence in English. *UCLA Working Papers in Phonetics*, 106, 1-33.

Cho, Y. Y. 1990. Syntax and Phrasing in Korean. In Inkelas, S., & Zec, D. (eds.) The Phonology-Syntax Connection: 47-62. Chicago: University of Chicago Press.

Christophe, A., Peperkamp, S., Pallier, C., Block, E., & Mehler, J. 2004. Phonological phrase boundaries constrain lexical access: I. Adult data. *Journal of Memory and Language*, 51, 523-547.

Cooper, W., & Paccia-Cooper, J. 1980. Syntax and Speech. Cambridge, MA: Harvard University Press.

Dart, S. 1987. An aerodynamic study of Korean stop consonants: Measurements and modeling. *Journal of Acoustic Society of America*, 81(1), 138-147.

Diehl, R.L., & Kluender, K.R. 1989a. On the objects of speech perception. *Ecological Psychology*, 1, 121-144.

Diehl, R.L., & Kluender, K.R. 1989b. Reply to commentators. *Ecological Psychology*, 1, 195-225.

Fougeron, C. 2001. Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, 29, 109-135.

Fougeron, C., & Keating, P. A. 1997. Articulatory strengthening at edges of prosodic domains. *Journal of the Acoustical Society of America*, 101(6), 3728-3740.

Ganong, W. 1980. Phonetic categorization in auditory word perception. *Journal of Experimental Psychology: Human Perception and Performance* 6. 110-125.

Guion, S. 1995. Word frequency effects among homonyms. Texas Linguistic Forum 35. University of Texas, Austin. 103-116.

Halle, M., & Stevens, K. N. 1971. A note on laryngeal features. Quarterly progress report (vol.101, pp. 198-122). Cambridge, MA: Research Laboratory of Electronics, MIT.

Han, J.-I. 1996. The Phonetic and Phonology of "Tense" and "Plain" Consonants in Korean. Ph. D. dissertation. Cornell University.

Han, M.S., & Weitzman, R.S. 1970. Acoustic features of Korean /P, T, K/, /p, t, k/ and /p^h, t^h, k^h/. *Phonetica*, 22, 112-28.

Hardcastle, W.J. 1973. Some observations on the Tense-Lax Distinction in Initial Stops in Korean. *Journal of Phonetics*, 1, 263-271.

Hay, J. F., Sato, M., Coren, A. E., Moran, C. L. & Diehl, R. L. 2006. Enhanced contrast for vowels in utterance focus: A cross-language study. *Journal of the Acoustical Society of America*, 119 (5), 3022-3032.

Jang, M. 2008. The Acoustic Characteristics of Aspiration Merger in Korean, presented at the Linguistic Society of America, Chicago, IL, 2008.

Jun, S.-A. 1993. The phonetics and phonology of Korean prosody. Doctoral dissertation. Ohio State University.

Jun, S.-A. 1998. The accentual phrase in the Korean prosodic hierarchy, Phonology, 15(2), 189-226.

Jun, S-A. 2000. "K-ToBI (Korean ToBI) labeling conventions: Version 3", *Speech Sciences*, 7, 143-169.

Jun, S.-A. & H. Lee. 1998. Phonetic and Phonological markers of Contrastive Focus in Korean, in *Proceedings of the 5th International Conference on Spoken Language Processing*, Vol. 4, p. 1295-1298, Sydney, Australia.

Jun, S.-A., Beckman, M.E, & Lee, H.-J. 1998. Fiberscopic evidence for the influence on vowel devoicing of the glottal configurations for Korean obstruents. *UCLA Working Papers in Phonetics* 96:43-68.

Kang, O. 1992. Korean Prosodic Phonology. Ph.D. dissertation. University of Washington.

Kang, K.-H., & Guion, G. S. 2008. Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *Journal of the Acoustical Society of America*, 124(6), 3909-3917.

Kagaya, R. 1974. A fiberscopic and acoustic study of Korean stops, affricates and fricatives. *Journal of Phonetics*, 2, 161-80.

Keating, P. A. 1995. Effect of prosodic position on /t, d/ tongue/palate contact, poster presented at the XIIIth International Congress of Phonetic Sciences, Stockholm, 1995 (unpublished).

Keating, P. A, Cho, T.-H., Fougeron, C., Hsu, C. 1999. Domain-initial articulatory strengthening in four languages. *UCLA Working papers in Linguistics* 97, 139-156.

Keating, P. A., Wright, R., & Zhang, J. 1999. Word-level asymmetries in consonant articulation. *UCLA Working Papers in Phonetics*, 97, 157-173.

Keating, P. A., Cho, T., Fougeron, C., & Hsu, C. 2003. Domain-initial strengthening in four languages. In J. Local, R. Ogden, & R. Temple (Eds.) *Papers in laboratory phonology 6: Phonetic interpretations* (pp. 145-163). Cambridge, UK: Cambridge University Press.

Kim, C.-W. 1965. On the autonomy of the tensity feature in stop classification. *Word*, 21, 339-359.

Kim, C.-W. 1970. A theory of aspiration. *Phonetica*, 21, 107-116.

Kim, M.-R., Beddor, P. S., & Horrocks, J. 2002. The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics*, 30 (1), 77-100.

Kim, H.-S. & Kiyoshi, H. & Shinji, M. 2005. Stroboscopic-cine MRI study of the phasing between the tongue and the larynx in the Korean three-way phonation contrast. *Journal of Phonetics*, *33*, 1-26.

Kim, H.-S. & Kiyoshi, H. & Shinji, M. 2009. Invariant articulatory bases of the features [tense] and [spread glottis] in Korean plosives: New stroboscopic cine-MRI data. *Journal of Phonetics* in publication.

Kim, S. 2001. The interaction between prosodic domain and segmental properties: domain initial strengthening of fricatives and post obstruent tensing rule in Korean. MA thesis, University of California, Los Angeles.

Kingston, J., & Diehl, R.L. 1994. Phonetic knowledge. *Language*,70, 419–454.

Krakow, R.A. 1989. The Articulatory Organization of Syllables: A kinematic analysis of labial and velic gestures, Ph.D. dissertation, Yale University.

Lindblom, B. 1990. Explaining phonetic variation: A sketch of the H & H theory. In W. Hardcastle and A. Marchal (eds.) Speech Production and Speech Modeling. Kluwer, Dordrecht, 403-439.

Lisker, L., & Abramson, A.S. 1964. Cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384-422.

Lombardi, L. 1991. Laryngeal features and laryngeal neutralization. Ph.D. dissertation, University of Massachusetts at Amherst.

Lombardi, L. 1995. Laryngeal features and privativity. *The Linguistic Review*, 12, 35-59.

Miller, G. A., & Nicely, P. E. 1955. An analysis of perceptual confusions among some English consonants, *Journal of the Acoustical Society of America*, 27, 338-352.

Munson, B. and Solomon, N. 2004. The effect of phonological density on vowel articulation. *Journal of Speech, Language and Hearing Research* 47. 1048-1058.

Nespor, M., & Vogel, I. 1986. Prosodic Phonology. Dordrecht: Foris publications.

Ohala, J. & Kawasaki, H. 1984. Prosodic phonology and phonetics, *Phonology Yearbook*, 1, 113-127.

Park, H.-S. 1999. The phonetic nature of the phonological contrast between the lenis and fortis fricatives in Korean. In *Proceedings of the 14th International Congress of Phonetic Science* (ICPhS 99), vol. 1, 424-427.

Pierrehumbert, J., & Beckman, M. 1988. Japanese Tone Structure. MIT, Cambridge, MA.

Pierrehumbert, J., & Talkin, D. 1992. Lenition of /h/ and glottal stops. In *Papers in Laboratory Phonology II: gesture, segment, prosody* (G. Docherty & D.R. Ladd, editors), 90-117. Cambridge, U. K.: Cambridge University Press.

Pluymaekers, M., Ernestus, M. and Baayen. R.H. 2005. Lexical frequency and acoustic reduction in spoken Dutch. *Journal of the Acoustical Society of America*, 118, 2561-2569.

Quené, H. 1992. Durational cues for word segmentation in Dutch. *Journal of Phonetics*, 20, 331-350.

Redford, M.A, & Diehl, R.L. 1999. The relative perceptual distinctiveness of initial and final consonants in CVC syllables. *Journal of the Acoustical Society of America*, 106(3), 1555-1565.

Selkirk, E. 1984. Phonology and Syntax: The Relation between Sound and Structure, MIT Press, Cambridge, MA, and London, England.

Selkirk, E. 1986. On derived domains in sentence phonology. Phonology Yearbook, 3, 371-405.

Shattuck-Hufnagel, S., & Turk, A. E. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistics Research*, 25, 193-247.

Stevens, K. N. 1999. Acoustic Phonetics. Cambridge, MA: MIT Press.

Stevens, K. N., Keyser, S.J., & Kawasaki, H. 1986. Toward a phonetic and phonological theory of redundant features. In *Invariance and Variability in Speech Processes* (J.S.Perkell & D. H. Klatt, editors), pp. 426-463. Hillsdale, NJ: Lawrence Erlbaum Associates.

Sussman, H. M., Bissell, N., Dalston, E., & Majors, T. 1997. An investigation of stop place articulation as a function of syllable position: a locus equation perspective, *Journal of the Acoustical Society of America*, 101, 2826-2838.

Thorsen, N. 1985. Intonation and text in Standard Danish. *Journal of the Acoustical Society of America*, 77, 1205-1216.

Umeda, N. 1982. F0 declination is situation dependent. *Journal of Phonetics*, 10, 279-290.

Vaissiére, J. 1988. Prediction of velum movement from phonological specifications, *Phonetica,*45, 122-139.

Van Coile, N. 1987. A model of phoneme durations based on the analysis of a read Dutch text, in Proc. Europ. Conf. Speech Techn., vol.2, edited by J. Laver and M. Jack Edinburgh, United Kingdom, 233-236.

Warren, R. 1970. Perceptual restoration of missing speech sounds. *Science*, 167, 392-393.

Wightman, C.W., Shattuck-Hufnage, S., Ostendorf, M., & Price, P. J. 1992. Segmental durations in the vicinity of prosodic phrase boundaries, *Journal of the Acoustical Society of America,* 91, 1707-1717.

Wright, R. 2003. Factors of lexical competition in vowel articulation, in Phonetic Interpretation: Papers in Laboratory Phonology VI, edited by J. Local, R. Ogden, and R. Temple (Cambridge University Press, Cambridge), 75-87.

Yoon, K. 1998. An acoustic analysis of Korean aspirated and tense alveolar fricatives using Multi-Speech. M.A. thesis. University of Kansas.

VITA

Mi Jang attended Namsan High School in Daegu, Korea on 1991. She entered Kyoungpook National University at 1994 and majored in Anthropology and Archaeology, and English language and literature as her second major. She received a Bachelor of Arts in 1998 and began graduate studies in English Language and Literature department at Kyungpook National University. After receiving Master of Art in 2000, she attended the Doctoral program in English Language and Literature department at Seoul National University on 2001. She was working as a researcher in the Brain Korea (BK) 21 project of Seoul National University from 2001 to 2002. From August, 2003 to present, she has attended the Graduate School in the Linguistic department at the University of Texas at Austin. She worked as a teaching assistant from 2008 to 2009.

Permanent Address: Tokyo-do Taitou-ku Ikenohata 2-1-44 Ikenohata
Nagatani mansion #1204, Japan, zip code: 110-0008

This dissertation was typed by the author.